



THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE
Ès sciences économiques

Délivré par l'Université Toulouse 1 Capitole (UT1 Capitole)
Cotutelle internationale avec l'Université de Tunis

Présentée et soutenue publiquement par

Inès MOUSSA

Le 30 novembre 2012

Proximité géographique, diffusion des connaissances et
innovation : une étude sur les départements français
métropolitains

Jury

Monsieur Mohamed Ayadi

Professeur à l'Institut Supérieur de Gestion de Tunis, directeur de thèse

Monsieur Olivier Brossard

Professeur à l'Institut d'Études Politiques de Toulouse, directeur de thèse

Madame Danielle Galliano

Directrice de Recherche à l'Institut National de Recherche Agronomique de Toulouse

Madame Zouhour Karray

*Maître de conférences HDR à la Faculté des Sciences Économiques et de Gestion de
Nabeul, rapporteur de thèse*

Madame Nadine Massard

Professeure à l'Université Jean Monnet de Saint-Étienne, rapporteur de thèse

Ecole doctorale Temps, Espaces, Sociétés, Cultures (TESC)
Laboratoire d'Étude et de Recherche sur l'Économie, les Politiques et les Systèmes Sociaux
(LEREPS)

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

Je dédie cette thèse à la mémoire de mon père.

Remerciements

Ce travail a été effectué dans le cadre d'une thèse en co-tutelle entre l'Institut Supérieur de Gestion de Tunis et l'Université de Toulouse 1 Capitole. Je remercie très cordialement les responsables de la coopération universitaire franco-tunisienne pour avoir financé mes séjours à Toulouse durant mes deux premières années de thèse.

Je tiens à exprimer ma profonde gratitude à mes deux directeurs de thèse qui ont su donner à ma recherche la direction qu'elle a suivie.

Mes vifs remerciements à Monsieur Mohamed Ayadi pour son aide inestimable, pour ses précieux conseils et pour la confiance qu'il m'a accordée dès le début de ma thèse. Ma sincère reconnaissance à Monsieur Olivier Brossard pour m'avoir fait partager ses nombreuses connaissances et pour m'avoir souvent donnée le courage d'avancer dans mes recherches en me remotivant lorsque j'en éprouvais le besoin. Ses conseils, ses remarques et sa lecture méticuleuse de toute la thèse m'ont permis de préciser mon propos et d'affiner mes développements. Sans lui, cette thèse n'aurait jamais été menée à terme.

Je souhaite ensuite remercier Madame Zouhour Karray et Madame Nadine Massard qui font l'honneur de rapporter cette thèse. Je remercie infiniment Madame Danielle Galliano pour avoir accepté de faire partie de ce jury.

Je me dois également de remercier Marie-Benoît Magrini, Mohamed Amara et Thibault Laurent pour m'avoir aidée à maîtriser les techniques de l'économétrie spatiale.

Je remercie de tout mon coeur tous les professeurs et doctorants de l'équipe du LEREPS que j'ai côtoyés pendant ces années de thèse. Merci plus particulièrement à Rachel Levy, Med Kechidi, Malika Hattab-Christmann, Jérôme Vicente, Nicolas Soulié, Pierre-Alexandre Balland, Huu Thanh Tam Nguyen et Luis Orozco pour leurs commentaires critiques sur mon travail.

J'adresse aussi mes remerciements à Simon Nadel, Geoffroy Labrousche, Achille Diendere, Christine Rocheteau, Catherine Baron, Xavier Marchand Tonel et Nouri Mbarek pour leurs relectures attentives.

Je salue chaleureusement mes compagnons d'aventure : Ayse, Liliana, Guihong, Susanna, Yasmina, Tania, Hélène, Mikaël, Joan, Aymeric, Tristan, Seb, Bertrand, Geoffroy, Arona, et bien d'autres. Leur sympathie et leur bonne humeur ont été un élément positif dans l'établissement de cette thèse. Je remercie également Anne-Marie et Sophie pour leur gentillesse.

Un grand merci à toute l'équipe de l'UAQUAP plus particulièrement Imen, Hanen, Olfa, Kamel, Essahbi et Mohieddine pour la plus belle amitié qui nous lie malgré la distance.

Enfin, mes remerciements vont à ma famille et spécialement ma mère Alia pour ses encouragements, son affection et pour tous les sacrifices qu'elle avait prodigués pour ma réussite. Je lui serai toujours redevable. Avoir porté ce travail à terme représente pour moi aujourd'hui la plus belle des récompenses.

Je remercie également mes grands parents, mon frère Ahmed, ma soeur Thouaiba, Mustapha et Imen pour leurs encouragements qui ont été très importants pour moi.

Que tous ceux qui m'ont aidée de près ou de loin dans l'élaboration de ce travail trouvent ici l'expression de ma sincère gratitude.

Sommaire

Introduction générale	11
1. Un éclairage théorique sur la question des frontières géographiques de la diffusion des connaissances	19
2. Les outils de mesure de la diffusion spatiale des connaissances : Quelle est la bonne approche empirique ?	71
3. Diffusion spatiale des connaissances en France : Une analyse exploratoire de données spatiales	135
4. De l'intérêt d'une analyse économétrique spatiale appliquée à des données de panel	173
Conclusion générale	205
Bibliographie	211
Annexes	243
Table des matières	259

Introduction générale

L'influence cruciale de l'innovation sur la croissance n'est plus à démontrer. L'économiste Joseph Schumpeter l'a largement soulignée dans ses différents écrits. Sous l'effet de la mondialisation et des progrès techniques apportés par les nouvelles technologies, les entreprises se sont appuyées plus fortement sur l'innovation afin de consolider leurs places sur le marché, et elles ont dû pour cela accroître leurs investissements en matière de recherche et développement (R&D).

En 2000, la commission européenne a fixé pour l'union européenne un objectif chiffré : consacrer 3% du PIB de l'Union Européenne aux dépenses en R&D¹, objectif dit de Lisbonne. Depuis sa mise en place, cette stratégie n'a pourtant pas réussi à dynamiser l'économie européenne. L'Europe des 27 est encore loin de l'objectif des 3% fixé en 2000 à Lisbonne. En effet, les dépenses en R&D sont aujourd'hui de 1,9% en moyenne selon les estimations d'Eurostat. Seules la Suisse (3,6%) et la Finlande (3,47%) dépassent ce seuil.

Afin d'aider les pays à atteindre ces 3%, l'Europe a mis en place différentes mesures permettant de dynamiser les dépenses en R&D à travers l'amélioration des partenariats entre les entreprises et les centres de recherche. Cependant ces mesures n'ont pas pleinement réussi à relever le défi de l'innovation. L'effort global de la R&D pour les entreprises plafonne actuellement autour de 0,8% du PIB, loin de l'objectif fixé il y a 10 ans, qui était de 2% du PIB.

La stratégie de Lisbonne n'a donc pas atteint les objectifs initialement fixés².

1. Il est estimé qu'une augmentation de 0,1% dans l'intensité de R&D fait croître le PIB par habitant de 0,3 à 0,4%.

2. Le parlement européen, dans sa résolution du 10 mars 2010 relative à l'Europe 2020, consacre un paragraphe aux causes de l'insuccès de la stratégie de Lisbonne, en citant :

- "*La faiblesse de la structure de gouvernance*";
- Le "*Manque de responsabilisation*";
- "*L'objectif général extrêmement complexe, fait de trop nombreux objectifs sectoriels*" et, enfin,
- "*Un excès d'ambition et un manque de précision, d'orientation claire et de transparence*".

Source : Rapport d'information déposé par le comité d'évaluation et de contrôle des politiques publiques sur l'évaluation des incidences de la Stratégie de Lisbonne sur l'économie française, enregistré à la Présidence de l'Assemblée nationale le 16 février 2012.

Pourtant, l'innovation c'est-à-dire l'invention de nouveaux produits, l'amélioration de la qualité ou la réduction du coût de production des biens et services existants, est le fruit de l'investissement en R&D. La R&D est le terme que nous utilisons pour désigner les efforts des scientifiques, ingénieurs, entrepreneurs et inventeurs qui développent de nouvelles connaissances et contribuent à leur diffusion.

Si l'investissement en R&D reste faible dans de nombreux pays européens, c'est peut-être que les enjeux en sont mal compris : avant d'inventer de nouveaux biens ou services, il faut accumuler des connaissances nouvelles, articuler de façon nouvelle des connaissances existantes, créer les conditions favorables à l'exploration de nouveaux savoirs. C'est précisément en cela que consiste la R&D, mais c'est aussi cela qui en fait un investissement à très long terme, dont la rentabilité est incertaine et dont l'appropriation privée est toujours imparfaite. Un enjeu tout à fait crucial et particulier de la R&D est sa dimension circulatoire : avant de donner lieu à des inventions, les connaissances nouvelles produites par les activités de R&D doivent circuler. C'est en effet en se déplaçant dans l'espace que ces connaissances vont d'une part s'enrichir et se transformer et d'autre part toucher les investisseurs potentiels susceptibles de les exploiter.

Nous pensons que la bonne circulation des connaissances accroît les succès de la R&D, ce qui, en retour, incite les entreprises et organisations de recherche publiques à faire encore plus de R&D. **C'est pourquoi il est essentiel de mieux comprendre la géographie des flux de connaissance et les facteurs qui contribuent à ce qu'elles circulent plus ou moins bien dans l'espace. Nous faisons de ce questionnement la problématique centrale de notre thèse.**

Dès lors, l'objectif de maximisation de l'output d'innovation n'apparaît plus uniquement dépendant du choix d'une quantité d'input de R&D (les 3% de l'objectif de Lisbonne) : il devient aussi dépendant d'un objectif d'optimisation de l'organisation géographique de la R&D et des flux de connaissance qui en résultent.

Par ailleurs, le phénomène de l'innovation n'est pas le fruit de la R&D d'un individu isolé dans la mesure où les connaissances des agents sont améliorées par celles d'autres agents. Ainsi, dès lors que ce processus résulte de l'implication de plusieurs personnes plus au moins distantes, l'espace a un rôle à jouer dans la dynamique d'accumulation et de création de connaissances.

Contexte de la recherche

Dans cette thèse, nous partons du principe que la diffusion des connaissances dans l'espace joue un rôle important dans le processus d'innovation. Une littérature

foisonnante a étudié la dimension spatiale des phénomènes économiques en général (Lucas, 1988 ; Romer, 1986 ; Walz, 1996) et de l'innovation en particulier (Caniëls, 2000 ; Carrincazeaux et al., 2001 ; Escribano et al., 2009). Au centre de ces approches théoriques et empiriques, se situe la notion d'externalités de connaissance localisées (LKS). Selon Jaffe (1986), les externalités de connaissance représentent des flux d'idées entre agents dont le coût ou le bénéfice social diffère du coût ou du bénéfice privé³. De tels phénomènes sont particulièrement associés à la production de connaissance. En effet, les propriétés économiques de la connaissance sont à l'origine de ces externalités. La raison tient au caractère difficilement appropriable de la connaissance (Arrow, 1962 ; Foray, 2000). Ces flux de connaissance sont représentés comme un des facteurs à l'origine de la localisation des outputs de l'innovation. Dans un tel contexte, les externalités de connaissance sont supposées géographiquement circonscrites, ce qui engendre des économies d'agglomération. La question du caractère spatialisé des externalités et de leur étendue géographique suscite donc un intérêt majeur.

Dès le début des années 1990, plusieurs travaux ont tenté de fournir une évaluation de l'étendue géographique des externalités de connaissance et de leur caractère spatialisé. Ces travaux peuvent être distingués en deux groupes.

Dans un premier groupe, sont utilisés des modèles de croissance endogène pour comprendre les différences de résultats économiques (croissance, productivité...). Les recherches empiriques inspirées de ces travaux évaluent donc l'effet des externalités de connaissance sur la croissance et la productivité (Coe et Helpman, 1995).

Le deuxième groupe est inspiré de l'économie géographique dans le sens où il essaie de fournir une mesure de la dimension géographique des externalités de connaissance. Il s'agit alors de mesurer l'étendue géographique de la diffusion des connaissances. Cette évaluation peut se faire de deux manières. La première est l'approche citation de brevets qui consiste à mesurer les externalités via leurs traces de diffusion (Jaffe, et al 1993 ; Jaffe et Trajtenberg, 1999 ; Almeida et Kogut, 1999). La deuxième manière consiste à utiliser la fonction de production de connaissance afin de quantifier leurs effets sur l'output de l'innovation plutôt que leurs traces dans les citations de brevets. Cette approche remonte aux travaux de Zvi Griliches (1979), elle fut reprise par Jaffe (1989) et développée par la suite par Anselin et al.(1997). Ces derniers ont construit des modèles basés sur l'interaction spatiale permettant de passer d'une approche "indirecte" qui déduit l'existence des externalités à une approche "directe" qui les mesure plus finement. L'interaction spatiale entre observations a été identifiée par des méthodes variables dans la littérature : par exemple en testant l'impact de la

3. La notion d'externalités de connaissance a été définie en détail dans le premier chapitre de notre travail.

R&D des départements limitrophes, ayant une frontière commune, sur l'innovation du département considéré (Autant-Bernard, 2001a,b), ou en mesurant la R&D des voisins dans des cercles concentriques (Bottazzi et Peri, 2003).

Dans la continuité de ces travaux, nous ambitionnons de contribuer empiriquement à mesurer l'étendue géographique des flux de connaissance tout en expliquant comment ces connaissances se diffusent et pourquoi l'espace peut être considéré comme un catalyseur de leur diffusion.

Problématique et thèse défendue

Les travaux effectués pour mesurer l'étendue géographique des flux de connaissance et pour révéler le rôle de la proximité géographique dans la production de l'innovation restent insuffisants (Carrincazeaux, 2011; Feldman, 2003). Même les approches les plus avancées, fondées sur l'estimation d'une fonction de production de connaissance, déduisent plus qu'elles ne démontrent la dimension locale des externalités. En effet, l'ambiguïté du concept d'externalité a rendu les interprétations empiriques une tâche difficile. Ainsi, l'intérêt et parfois même l'existence des externalités de connaissance ont été remis en cause (Breschi et Lissoni, 2003). Partant de ce constat, il devient essentiel d'utiliser une approche empirique plus approfondie et plus directe pour mesurer d'une manière plus fiable l'ampleur de la diffusion des connaissances dans l'espace ainsi que leurs effets sur la production d'innovation.

Nous ne cherchons donc pas à savoir si les connaissances diffusées sont des externalités ou non, mais plutôt à déterminer l'étendue géographique des connaissances diffusées et aussi à cerner les facteurs qui font que le processus de diffusion soit localisé ou distant.

Comme dans beaucoup d'autres travaux sur le sujet, nous emploierons abusivement le terme "externalités" pour traduire "*spillovers*" même lorsqu'il s'agira de désigner de simples flux de connaissance circulant dans l'espace et dont nous ne pouvons pas vérifier qu'il s'agit de pures externalités. De cette problématique globale découlent d'autres questions :

- Les connaissances qui se diffusent dans l'espace sont-elles toutes localisées (*Knowledge Localised, KL*)? Existe-il des connaissances délocalisées qui se diffusent au-delà de l'espace géographique proche (*Knowledge Distant, KD*)? Si oui, comment se diffusent-elles et quel est le rôle des réseaux d'innovation et des *clusters* dans leur diffusion?
- Quelle est la relation entre la nature de la connaissance et l'ampleur de sa diffusion? Est-il vrai que les connaissances tacites sont plus localisées dans

l'espace que les connaissances codifiées? Peut-on mesurer d'une manière plus directe l'ampleur de l'étendue géographique de ces deux types de connaissance?

Pour ce faire, nous présentons trois hypothèses de recherche que nous espérons tester à travers cette étude :

Hypothèse 1 : La R&D interne intègre à la fois des connaissances tacites et des connaissances codifiées alors que la R&D externe ne véhicule pratiquement que des connaissances codifiées. Par conséquent, en introduisant ces deux mesures différentes dans notre modèle, nous pourrions distinguer les limites géographiques de chacune des connaissances tacites et codifiées.

Hypothèse 2 : les clusters dotés des réseaux de diffusion des connaissances les plus variés et les plus efficaces ont une meilleure productivité à l'innovation car ils tirent partie des *spillovers* de connaissance plus intenses émanant de ces réseaux. Dans les clusters high-tech, ces différents canaux sont complémentaires et ont des effets cumulatifs plutôt que substituables.

Hypothèse 3 : Puisque les industries similaires ont tendance à choisir des localisations similaires, la recherche de connaissances dans d'autres secteurs peut nécessiter des interactions à distance avec les détenteurs de connaissances situés dans des régions voisines. Il importe donc pour l'innovation régionale que les régions voisines soient diversifiées plutôt que hautement spécialisées.

En plus de ces hypothèses que nous testons empiriquement dans les chapitres 2, 3 et 4, nous tentons dans le chapitre 1 de clarifier la notion particulière d'externalités de connaissance et les notions plus générales de diffusion des connaissances ou de flux de connaissance. Nous y discutons aussi des déterminants de l'étendue géographique de ces flux et de leur influence sur la localisation et la productivité de l'innovation.

Architecture de la thèse

Pour répondre à nos questions de recherche formulées ci-dessus, ce travail se divise en quatre chapitres. Notre premier chapitre est consacré à présenter les théories liées à la diffusion des connaissances. Ce chapitre permet de poser les fondements théoriques ainsi que pour comprendre la suite de notre développement empirique. Il est suivi de trois chapitres empiriques qui présentent notre première méthode de mesure des externalités de connaissance (chapitre 2) et notre deuxième méthode basée sur les outils de l'économétrie spatiale (chapitre 3 et chapitre 4) :

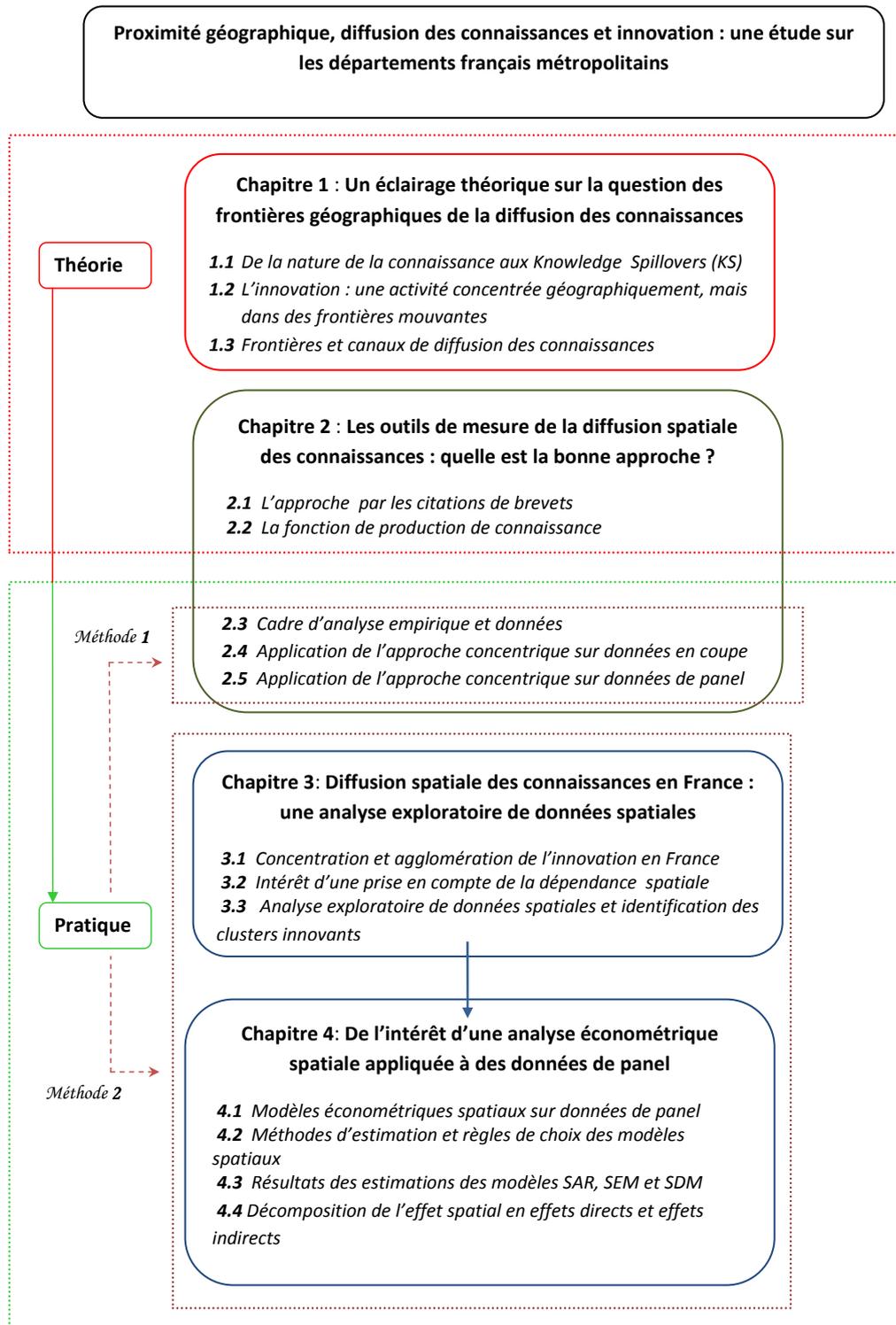


Figure 1 – Architecture de la thèse

Le premier chapitre a pour objectif de revenir sur les concepts de la connaissance et des externalités de connaissance tout en synthétisant la littérature sur la relation connaissance-externalités-innovation. Après une présentation des propriétés économiques de la connaissance qui sont à l'origine des externalités (section 1), nous consacrons la section 2 à l'étude des effets des externalités de connaissance localisées et de l'espace dans la production d'innovation. Dans la section 3, nous montrons l'importance des canaux de transmission des connaissances dans le processus d'innovation. Ensuite, nous précisons comment les réseaux d'innovation permettent la diffusion des connaissances délocalisées.

Nous proposons dans cette thèse d'analyser empiriquement l'impact de l'espace et de la diffusion des connaissances sur l'innovation. Pour cela, nous proposons dans **le deuxième chapitre** une première approche empirique par cercles concentriques afin de mesurer l'étendue géographique des connaissances diffusées. Dans un premier temps, nous présentons les principaux modèles empiriques étudiant la relation entre innovation et externalités de connaissance. Ces travaux sont résumés en deux approches : l'approche qui mesure les flux de connaissance par les citations de brevets et l'approche par les fonctions de production de connaissance qui mesure plus directement les externalités à travers la prise en compte des interactions entre unités voisines. Dans un deuxième temps, nous tentons d'appliquer l'approche par cercles concentriques de Bottazzi et Peri (2003) sur un échantillon de données françaises. A l'aide de données relatives à 86 régions européennes, Bottazzi et Peri (2003) mesurent l'étendue géographique des externalités de connaissance dans différents niveaux géographiques, dans la région elle-même, dans les régions situées à moins de 300 kilomètres, dans les régions situées entre 300 et 600 kilomètres, entre 600 et 900 kilomètres, entre 900 et 1300 kilomètres et entre 1300 et 2000 kilomètres. Les mesures effectuées pour ce cas sur des données en coupe couvrant la période 1977-1995 montrent que les effets des externalités se dissipent avec la distance et que seulement la R&D mesurée dans l'intervalle $[0, 300km[$ influence l'innovation de la région considérée.

Partant de l'approche par cercles concentriques de Bottazzi et Peri (2003), nous exploitons les données relatives aux 94 départements français métropolitains pour mesurer l'étendue géographique des connaissances diffusées. Au départ, nous exploitons comme Bottazzi et Peri (2003) des données en coupe transversale, mais nous appliquons ensuite la même méthode sur des données de panel. Les résultats obtenus sont divergents et instables et cela en cause nous conduit à utiliser d'autres méthodes.

C'est pourquoi, nous utilisons dans **le troisième chapitre** les techniques de l'économétrie spatiale afin de mesurer de manière plus fine l'ampleur des effets d'interdépendance dans l'espace. Les techniques de l'autocorrélation spatiale exploitées

dans le chapitre 3 de notre travail ont pour objectif d'analyser la structure de concentration géographique des activités innovantes. Au-delà de la simple observation de la concentration géographique, ce chapitre a pour objectif de détecter l'existence de l'autocorrélation spatiale entre les unités spatiales tout en apportant de nouveaux éclairages sur la nature de la causalité entre les observations spatiales. Les outils mobilisés dans ce chapitre sont : le diagramme de Moran, les statistiques de LISA et les cartes de significativité de LISA qui sont une combinaison des deux derniers indicateurs. Le diagramme de Moran permet seulement de repérer les différents régimes spatiaux et de déterminer les localisations qui correspondent ou non au schéma global d'autocorrélation spatiale. Les statistiques de LISA nous précisent les niveaux de significativité de ces associations spatiales. Finalement, l'utilisation conjointe des statistiques de LISA et du diagramme de Moran nous permettent de déterminer les clusters innovants à autocorrélation spatiale significative. En revanche, même si ce chapitre nous a permis de détecter les clusters innovants, ceci n'est qu'une analyse exploratoire et les résultats trouvés dans ce chapitre ne nous permettent pas de mesurer l'ampleur de la diffusion des connaissances. Pour ce faire, nous poursuivons notre analyse spatiale dans le chapitre 4.

Le quatrième chapitre mesure d'une manière plus directe l'étendue spatiale des connaissances diffusées. Dans une première partie, nous proposons de tester 3 modèles spatiaux : un modèle spatial autorégressif (SAR), un modèle avec autocorrélation des erreurs (SEM) et un modèle régressif croisé (SDM). Notre objectif est de déterminer le modèle spatial le plus approprié pour nos estimations et de mesurer l'ampleur des dépendances spatiales au niveau de nos variables. Nos résultats montrent que l'effet spatial existe mais d'une ampleur faible et que les élasticités des dépenses en R&D (interne et externe) présentent des signes attendus quel que soit le modèle utilisé.

Dans la deuxième partie de ce chapitre, nous décomposons les effets spatiaux en effets directs et effets indirects. En effet, cette décomposition nous permet de déterminer l'ampleur de l'effet des externalités, mesurées par l'effet indirect, dans la production d'innovation (LeSage et Pace, 2009). L'ampleur de l'effet des externalités de connaissance diminue avec la distance. L'effet direct, quant à lui, tient compte de l'effet de rétroaction qui est le résultat de l'augmentation de l'innovation des départements voisins à un département i grâce à l'investissement de ce dernier dans les dépenses en R&D.

Pour finir, notre conclusion générale synthétise les résultats obtenus et propose quelques pistes de recherche futures.

Chapitre 1

Un éclairage théorique sur la question des frontières géographiques de la diffusion des connaissances

1.1. Introduction

La relation entre connaissance et innovation a pris une place prépondérante dans la littérature. Dès 1870, Alfred Marshall introduit le concept de connaissance dans sa théorie des districts industriels via ce qu'il nomme "*atmosphère*" (Marshall, 1919). Il contribue ainsi à l'émergence de l'approche de l'économie fondée sur la connaissance au sein de laquelle celle-ci est considérée comme le vecteur de l'innovation (Foray, 2000 ; OCDE, 2000). En effet, la diffusion des connaissances et l'apprentissage sont à l'origine d'une abondante littérature dans la mesure où "*knowledge is the most important strategic resource and learning the most important process*" (Lundvall, 1992 ; Morgan, 1997). Ainsi, l'échange d'idées revêt une dimension essentielle à l'accumulation du savoir, à sa mobilisation rapide et efficace, et à l'amélioration de la capacité cognitive (Mohnen, 1990 ; Nadiri, 1993 ; Maïresse et Sassenou, 1991).

En parallèle aux modèles de l'économie fondée sur la connaissance, d'autres études théoriques et empiriques sont apparues pour montrer que la diffusion des connaissances est limitée géographiquement : l'espace joue un rôle central dans leurs accumulations et leurs créations. Plus précisément, l'importance de ce phénomène dans la production de connaissance et d'innovation a été prise en compte avec l'apparition de la nouvelle économie géographique (NEG), qui a pour objectif l'étude du processus d'agglomération

spatiale des activités économiques au cours de ces deux dernières décennies (Palivos et Wang, 1993 ; Englmann et Walz, 1995 ; Kubo, 1995 ; Walz, 1996 ; Martin et Ottaviano, 1996 et 1999).

Le croisement de l'économie fondée sur la connaissance et de l'économie géographique a favorisé l'émergence de ce qu'il est aujourd'hui convenu d'appeler la "*géographie de l'innovation*" (Feldman, 2003 ; Carrincazeaux, 2011). Ces études sont apparues depuis la fin des années 90 afin de décrire la structure géographique des activités innovantes au sein des pays et des régions. Elles visent généralement à expliquer les phénomènes d'interdépendance entre firmes ou régions liés au fonctionnement de la connaissance et de l'innovation (Puga, 1998 ; Amati, 1998 ; OST, 1998 ; Paci et Usai 2000 ; Vertova, 2004 ; Autant-Bernard et al. 2008, Escribano et al, 2009).

L'objectif central de ces travaux est d'étudier le rôle de la proximité géographique dans la production d'innovation. Au centre de ces approches théoriques et empiriques, se situe la notion d'externalités de connaissance localisées (LKS) qui constitue le ferment principal de regroupement des firmes au niveau régional ou local (Massard et Torre, 2004). L'effet positif de la proximité géographique sur l'innovation résultait en effet pour l'essentiel de l'existence d'une limitation à la diffusion des externalités de connaissance dans l'espace (Autant-Bernard et al., 2008). Le recours aux externalités de connaissance localisées devient donc incontournable dès lors qu'il s'agit d'analyser la polarisation des activités innovantes.

Toutefois, d'autres analyses plus récentes affirment qu'aucune raison ne justifie qu'un processus d'apprentissage ne soit limité territorialement (O'Brien, 1992 ; Amin et Cohendet, 2004 ; Giuliani et Bell, 2005 ; Escribano et al., 2009). Ces approches remettent en cause l'importance des travaux des géographes pour deux raisons. La première raison relève de la prise en compte d'autres types de proximités non géographiques, notamment la proximité sociale, organisationnelle, institutionnelle ou cognitive (Boschma, 2005). Ceci a, d'une part, relativisé le poids de la proximité géographique et d'autre part, facilité le partage des connaissances (Boschma et Frenken, 2010 ; Balland, 2011).

La deuxième raison concerne le concept "*d'externalités de connaissance localisées*". En effet, ces nouvelles approches remettent en cause l'importance et parfois même l'existence des externalités de connaissance localisées (Breschi et Lissoni, 2001). Il y a eu donc un consensus sur le fait que les externalités de connaissance ne sont ni entièrement localisées, ni complètement libres ou diffusables à n'importe quelle distance et entre tout type d'agents ou lieux (Breschi et Lissoni, 2001 ; Amin et Cohendet, 2004 ; Boschma, 2005 ; Rallet et Torre, 2007). En d'autres termes, les approches soutenues par

ces économistes présentent des insuffisances et des limites dans la mesure où elles ont utilisé excessivement la notion "*d'externalités de connaissance*".

Au delà de ces limites, les activités innovantes restent localisées. A l'ère de la mondialisation et des nouvelles technologies, nous pouvons nous attendre à une dispersion de ces activités. Pourtant, la réalité est toute autre. Partant de ce constat, il nous semble utile donc de comprendre comment se diffusent les connaissances dans l'espace et pourquoi ce dernier reste un catalyseur à leurs diffusions. Faut-il encore considérer que les connaissances qui se diffusent soient toutes localisées? Ou bien, existe-il des connaissances délocalisées qui se déploient indépendamment de l'espace géographique considéré?

Partant de ces questionnements, l'objectif de ce chapitre est de présenter une revue de la littérature qui permet de cerner les principaux facteurs qui font que le processus de diffusion des connaissances soit localisé ou plutôt distant. Plus précisément, notre ambition ne consiste pas à savoir si les connaissances diffusées sont des externalités de connaissance ou simplement des flux de connaissance, mais à étudier l'étendue géographique (localisée ou distante) de ces connaissances, tout en précisant leurs modes de transmissions. Nous dépassons la simple question récurrente de l'espace pertinent au déploiement des externalités de connaissance pour étudier l'importance des canaux de leurs transmissions afin d'appréhender leur nature localisée, largement défendue par la littérature sur la géographie d'innovation. Dans un tel contexte, il serait important de passer en revue et de synthétiser la littérature sur les relations "*connaissance-externalités*" et "*connaissances-innovation*". C'est l'objectif du présent chapitre.

Dans section 2, nous cherchons à présenter les propriétés de la connaissance qui sont à l'origine de ces externalités. Au delà de sa nature tacite/codifiée, nous présentons l'aspect collectif et composite qui fait que la connaissance peut être diffusée plus loin dans l'espace. Une fois que ces externalités sont définies et distinguées, notamment les externalités de connaissance localisées (LKS), la section 3 sera dédiée à l'analyse des effets de la connaissance, de l'espace et des LKS tels qu'ils sont présentés dans la littérature sur la production d'innovation. Dans la section 4, nous essayons de vérifier théoriquement un constat largement défendu par les géographes : l'espace qui est facteur primordial dans le transfert des connaissances et dont les LKS se situent au centre de leurs analyses. Plus précisément, l'objectif de cette section est de montrer dans un premier temps que si l'espace compte dans le processus d'innovation, c'est parce qu'il existe d'autres facteurs tels que les canaux de transmission et de transfert des connaissances, en plus des LKS. Dans un deuxième temps, nous nous intéressons aux réseaux d'innovation afin de montrer comment ces derniers permettent la diffusion des

connaissances délocalisées (DKS), d'où l'aspect multidimensionnel de la connaissance. Enfin, la section 5 conclut le chapitre.

1.2. De la nature de la connaissance aux Knowledge Spillovers (KS)

Dans la théorie traditionnelle, la notion "d'externalités" a été basée sur l'assimilation de la connaissance à de l'information pure (Cohendet et Llerena, 1999). Dès qu'une connaissance est produite, elle est supposée s'échapper (par exemple à travers la mobilité des chercheurs) de sa source créatrice et devenir automatiquement utilisable au moindre coût par tous les autres acteurs de l'économie, réduisant de ce fait les incitations à produire cette connaissance.

Toutefois, cette approche a le défaut d'ignorer ce que pourrait apporter la distinction entre connaissance et information puisqu'elle néglige les thèmes de l'apprentissage, de la cognition et de la création intellectuelle (Foray, 2000). Partant de ce constat, les travaux de l'économie de la connaissance ont distingué la connaissance de l'information : la connaissance n'est plus considérée comme un stock d'informations mais une structure qui peut être transformée ou non par l'arrivée de messages (Stiglitz, 2000 et 2002). Ainsi, la connaissance a quelques chose de plus que l'information : elle offre la capacité "d'engendrer, extrapoler et inférer de nouvelles connaissances et informations" (Steinmueller, 1999, p372).

Pour Foray (2000), "la connaissance est d'abord fondamentalement une capacité d'apprentissage et une capacité cognitive, tandis que l'information reste un ensemble de données formatées et structurées, d'une certaine façon inertes ou inactives, ne pouvant pas elles-mêmes engendrer de nouvelles informations" Foray (2000, p9). Une personne qui possède des connaissances dans certains domaines, théoriques ou pratiques, est capable de produire à la fois de nouvelles connaissances et de nouvelles informations relatives à ce domaine. Cette distinction pourrait limiter la diffusion supposée automatique des externalités.

Néanmoins, la connaissance présente des caractéristiques "complexes" notamment l'imparfaite appropriabilité de ce bien qui engendre "des externalités" même dans le cas où la connaissance n'est pas assimilée à l'information. Il est important alors de s'interroger sur la manière dont la littérature considère le lien entre connaissance et externalités de connaissance. Ainsi, nous présentons dans ce qui suit les caractéristiques de la connaissance qui expliquent les possibilités des importantes fuites de cette dernière, ainsi que les différents concepts utilisés pour définir ce phénomène de débordement.

1.2.1. Propriétés économiques de la connaissance et externalités

La connaissance est un bien économique particulier, possédant des propriétés différentes de l'information et de celles des biens conventionnels (Arrow, 1962). En effet, la principale caractéristique de la connaissance est qu'elle a les caractéristiques d'un bien public. Elle est difficilement contrôlable, non rivale et cumulative (§1.2.1.1). Elle se diffuse et est utilisée par des tiers sans compensation pour les auteurs. Il y a donc des externalités positives (§1.2.1.2).

1.2.1.1. Propriétés économiques de la connaissance

La connaissance est *un bien non excluable*, c'est-à-dire qu'il est difficile de rendre ce bien exclusif, de le contrôler de façon privée : "*if nature has made any one thing less susceptible than all others of exclusive property, it is the action of the thinking power called an idea [...] he who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me*"¹(Jefferson, 1813, p.333).

Il en résulte donc des retombées involontaires, ce que la littérature économique appelle "*externalités*". Autrement dit, la non-appropriabilité des connaissances par l'entreprise la pousse à émettre des effets externes qui peuvent être utilisés librement par des agents n'appartenant pas à l'organisation émettrice de la connaissance, sans qu'ils n'aient à supporter les coûts de la recherche et du développement.

Les agents peuvent alors utiliser sans coût et quasiment sans limite des connaissances produites par d'autres agents. Par ailleurs, la connaissance est *un bien non rival*² en raison de son caractère inépuisable (Foray, 2000). En d'autres termes, la connaissance peut être utilisée plusieurs fois, par un nombre quelconque d'agents, et cela simultanément et sans se détériorer. Cette propriété de non-rivalité possède deux dimensions. Premièrement, un agent peut recourir une infinité de fois à la même

1. Lettre de Thomas Jefferson à Isaac McPherson, 13 août 1813, dans The Writings of Thomas Jefferson, Albert Ellery Bergh, 1907, p. 333-334, citée dans James Boyle, The Second Enclosure Movement and the Construction of the Public Domain, Deadalus, printemps 2002, p. 53.

2. L'implication de la propriété de non-rivalité sur les coûts et les prix est importante dans la mesure où la circulation de la connaissance se fait à un coût direct nul ou presque. La conséquence directe de cette propriété, en termes économiques, est que le coût marginal de l'utilisation d'une connaissance est nul. Puisque, une fois qu'une connaissance a été réalisée, ceci entraîne un coût fixe mais le coût de sa reproduction est nul, ce qui constitue une forte incitation à l'imitation. En effet, l'imitateur, contrairement à l'inventeur initial, n'encourt que le coût de production directe du bien, et non le coût de l'invention. Aussi, l'économie ne peut se conformer aux règles de fixation des prix sur la base des coûts marginaux car l'usage de la connaissance existante serait alors gratuit et il deviendrait impossible de compenser financièrement le fait qu'une connaissance soit utilisée un grand nombre de fois.

connaissance, sans qu'il lui en coûte, pour reproduire une action. Deuxièmement, une infinité d'agents peut utiliser la même connaissance sans que personne n'en soit privé : l'utilisation d'une idée par un agent n'empêche pas d'autres agents de l'utiliser simultanément : les connaissances sont donc non rivales dans leurs utilisations (Romer, 1993).

En plus de son caractère de non-exclusivité et de non-rivalité, la connaissance est aussi *un bien cumulatif*. Une connaissance est le facteur principal de la production de nouvelles connaissances et de nouvelles idées. Grâce aux propriétés de non-exclusivité et non-rivalité dans l'usage, la connaissance produit elle-même les conditions de l'amplification de ses effets en augmentant la production du savoir. Un théorème mathématique par exemple, s'enrichit à l'usage, prouvant sa fécondité et élargissant le champ de ses applications (Callon, 1999). Ainsi, la connaissance ne constitue pas seulement un bien de consommation mais aussi un bien de production, susceptible d'engendrer de nouveaux biens qui seront eux-mêmes infiniment utilisables. Le cumul des trois propriétés de non-contrôlabilité, non-rivalité et cumulativité accroît fortement le rendement social de l'activité de recherche et d'innovation, et inscrit celle-ci comme un fondement essentiel de la croissance.

La connaissance présente donc certaines caractéristiques qui expliquent le fait que sa production n'est pas le fruit de l'investissement d'un seul individu isolé mais plutôt d'un ensemble d'agents. Il découle de cette caractéristique de la connaissance des retombées positives de l'innovation qui sont bénéfiques à la société même si elles sont à l'origine du problème du bien public lié au fait que ces externalités agissent positivement sur le niveau de bien-être social, mais négativement sur celui des agents privés.

Encadré 1.1 : dilemme de la connaissance

Deux problèmes découlent de la caractéristique du bien public de la connaissance.

Le premier a trait aux retombées positives de l'innovation, c'est-à-dire au fait que la rentabilité sociale de l'innovation est généralement plus élevée que la rentabilité privée : un inventeur doit donc prévoir de recevoir moins que le rendement social de son invention. Il s'agit là du cas typique de défaut d'incitation et "*du problème du bien public*", qui conduisent à un niveau d'investissement privé insuffisant du point de vue de la société [Foray (2000)]. C'est un problème général décrit par Pigou dès 1932, et repris par Arrow en 1962a pour ce qui concerne le domaine de la recherche et l'innovation.

Le second problème est une autre facette du premier, dans la mesure où le savoir ne peut faire l'objet d'une appropriation. Dans ce cas, la firme ne peut retirer tous les bénéfices générés par son innovation, ce qui peut amoindrir l'incitation à investir dans des activités d'innovation. De ceci ressort ce que Foray appelle le dilemme de la connaissance. Le dilemme peut alors se résumer ainsi : "*seule l'anticipation d'un prix positif de l'usage garantira l'allocation de ressources pour la création; mais seul un prix nul garantira un usage efficient de la connaissance, une fois celle-ci produite*" (Foray, 2000 ; p.66). Si on augmente les incitations (en permettant aux firmes de s'approprier les connaissances) alors on diminue la diffusion, mais si on augmente la diffusion alors on décroît les incitations). Ce dilemme a pour objectif d'une part d'assurer à l'échelle de la société un usage efficient de la connaissance et d'autre part, une fois que cette connaissance est produite, de fournir une motivation idéale au producteur privé. Ainsi, contrairement à une externalité négative, il n'est pas possible de corriger une externalité positive, engendrée par un bien non rival et cumulatif. D'ailleurs, le problème n'est pas de réduire celle-ci puisqu'elle est positive. Par contre, dans le cas d'une externalité négative (bruit, pollution,...), nous pouvons agir sur l'émetteur en exigeant une correction à la source à travers la taxation par exemple.

1.2.1.2. Définitions et source des externalités de connaissance

Nous examinons dans cette partie les différentes définitions employées dans la littérature pour désigner ce phénomène d'externalités avant d'expliquer son origine et pourquoi ce phénomène a été largement lié à l'espace.

Définitions : En anglais, "*Spillovers*" signifie littéralement "*a side effect arising from or as if from an unpredicted source*"³, ce qui veut dire un effet secondaire découlant d'une source imprévue.

Les économistes ont utilisé le terme "*spillovers*" pour désigner les idées qui traversent volontairement ou non les frontières de la firme ou de la région innovatrice initiale, ce qui crée un exemple de "*positive externality*" (Marshall, 1920). Il en ressort de cette

3. The American Heritage dictionary of the English Language, Fourth Edition (2005).

définition littérale que la notion d'"*externality*"⁴ peut désigner aussi des externalités négatives contrairement à la notion de "*spillovers*" qui est utilisée que pour nommer les externalités positives. Néanmoins, les travaux qui étudient ce phénomène utilisent généralement "*externalités*" pour parler des "*externalités positives*" même si cette distinction n'est pas mentionnée dans la plupart des travaux économiques traitant ce phénomène et son impact sur l'innovation.

Après les *spillovers* ou les externalités positives, un autre concept est apparu, celui de "*Knowledge Spillovers*" ou "*externalités positives de connaissance*". Elles se produisent lorsqu'une entreprise s'approprie des connaissances provenant d'une autre firme, université ou organisme gouvernemental, sans payer pour cet avantage. Par conséquent, de tels phénomènes sont associés à la production de connaissances (Jaffe, 1986 ; 1989). Dans la plupart des études théoriques et empiriques sur les *spillovers*, les termes externalités de connaissance (Anselin et al., 2000 ; Feldman, 2003), Knowledge spillovers (Jaffe et al., 1993 ; Breschi et Lissoni, 2001) ou flux de connaissance (Peri, 2005 ; Agrawal et al., 2006) sont indifféremment utilisés pour désigner les externalités positives de connaissances. Ces externalités représentent le bénéfice de connaissances perçu par une firme ou une personne qui n'est pas responsable de l'investissement lié à la création de cette connaissance (Almeida et Kogut, 1999).

Les externalités de connaissance, appelées aussi technologiques⁵ (Griliches, 1979) peuvent désigner le fait que les entreprises peuvent acquérir des connaissances créées par d'autres sans avoir à payer pour cette connaissance dans une transaction de marché, et que les créateurs (ou les propriétaires actuels) n'ont aucun recours efficace, en vertu des lois en vigueur, si d'autres entreprises utilisent la connaissance ainsi acquise (Grossman et Helpman, 1992). Dans ce cas, le bénéficiaire des *spillovers* peut utiliser les nouvelles connaissances pour imiter les produits ou les procédés de l'innovateur ou bien il peut utiliser les connaissances en tant qu'input à son processus de recherche qui, à son tour, conduit à d'autres nouvelles technologies (Jaffe et al., 1993 ; Feldman, 1994 ; Audretsch et Feldman, 1996).

4. En général, une externalité est un effet positif ou négatif, qu'un agent économique produit sur un ou plusieurs autres agents. Nous parlons d'externalités négatives quand un acteur est défavorisé par l'action de tiers sans qu'il en soit compensé. La pollution est l'exemple le plus typique d'externalité négative dans la mesure où une usine qui pollue son environnement en rejetant des déchets, inflige une nuisance aux habitants de la région. Dans the American Heritage dictionary, externality signifie '*A cost or benefit that affects people other than those involved in the economic activity that produced it and that is not reflected in prices : pollution and other negative externalities.*'

5. Dans certaines études, la terminologie "*externalités technologiques*" est utilisée pour distinguer les externalités produites grâce à la proximité technologique entre entreprises ou régions, des "*externalités géographiques*" assimilées aux investissements en connaissances et cumulées à travers une certaine proximité géographique entre ces mêmes entreprises ou régions. Pour désigner les deux types de connaissances : géographique et technologique, nous employons la notion "*externalités de connaissance*".

D'autres notions ont été aussi utilisées dans la littérature pour désigner ces externalités comme par exemple les externalités de R&D. Ainsi, les études théoriques et empiriques ont montré de plus en plus que la R&D d'une entreprise ou d'une région peut avoir des effets positifs sur la productivité d'une autre entreprise ou région : une innovation dans les inputs permet des réductions des coûts et une innovation dans les produits par exemple favorise des expansions de marchés.

Par ailleurs, Griliches (1979) a précisé qu'"il y a deux notions distinctes d'externalités de R&D, souvent confondues dans la littérature". Il est important de noter que les externalités qui découlent de l'investissement en R&D ne sont pas seulement des externalités de connaissance, mais aussi des externalités pécuniaires⁶. Cette remarque rejoint celle de Breschi et Lissoni (2001) qui précisent que l'effet positif de l'espace sur l'innovation n'est pas seulement dû aux externalités de connaissance mais à d'autres phénomènes notamment les externalités pécuniaires.

En revanche, que nous parlions d'externalités de connaissance, externalités de R&D ou externalités technologiques, un large consensus reconnaît à ce phénomène de débordement une influence cruciale sur la croissance et l'innovation. Les résultats des études sur les *spillovers* montrent que les connaissances génèrent non seulement des *spillovers*, mais aussi que ces derniers ont tendance à être géographiquement délimités dans la région où les connaissances ont été créées et que l'étendue spatiale de ces externalités de connaissance est bornée.

Les activités de R&D : une source d'externalités de connaissance De même, Foray (2000) distingue deux façons de produire la connaissance qui à son tour constitue une source d'externalités : tout d'abord, les activités de recherche et développement (R&D) et ensuite l'apprentissage par la pratique⁷. Cette dernière est liée aux activités de R&D dans la mesure où aucune activité d'apprentissage ne pourrait avoir lieu sans un investissement initial en R&D. Pour cette raison, l'investissement en R&D constitue selon Foray (2000) la source la plus importante des externalités de connaissance.

En se référant comme plusieurs à la définition du rapport de Frascati (2002), l'organisation de coopération et de développement économique (OCDE) a défini la recherche et développement (R&D) comme un ensemble d'activités entreprises de façon

6. Si les externalités de connaissance proviennent de l'appropriation imparfaite de la connaissance, les externalités pécuniaires se produisent lorsque les prix d'intrants ne tiennent pas compte des améliorations de qualité réalisées via l'innovation, à cause soit de l'imperfection des prix, soit de l'asymétrie de l'information ou des coûts de transaction. Dans ce cas, le prix d'un bien intermédiaire n'est pas vendu à son coût marginal et la firme qui acquiert le bien bénéficie donc d'une externalité dans le sens où elle ne paye pas la totalité du prix. Nous parlons alors d'externalités pécuniaires.

7. L'apprentissage peut être soit fondé sur la répétition et le développement de l'expertise dans la mesure où la répétition d'une tâche permet de mieux l'accomplir ou bien fondé sur les expériences afin d'identifier une meilleure stratégie opérationnelle (Alder et Clark, 1991).

systématique en vue d'accroître des connaissances. Cette somme de connaissances⁸ sera utilisée pour de nouvelles applications, dans la réalisation d'une initiative, d'un projet ou autre. La R&D représente donc les efforts des scientifiques, ingénieurs, entrepreneurs ou inventeurs qui développent de nouvelles connaissances et accumulent les retombées du succès. Plusieurs modèles attribuent la croissance de la productivité aux activités de R&D (Romer, 1990 ; Aghion et Howitt, 1990). Si d'une entité à une autre, les ressources consacrées à la R&D diffèrent, ces activités tentent de faire progresser la technologie. La première conséquence de la R&D est la production de nouveaux produits ou procédés. Une fois celle-ci réalisée, ces nouveautés ont tendance à se diffuser à d'autres entités. Il s'ensuit alors des retombées privées et sociales qui relèvent en général le niveau de vie.

L'activité de R&D englobe trois sous-activités : la recherche fondamentale, la recherche appliquée et le développement expérimental. La recherche fondamentale est une recherche théorique sans objectif d'application ou d'utilisation postérieure. La recherche appliquée, contrairement à la recherche fondamentale qui ne cherche que la production de nouvelles connaissances générales, a pour objectif de regrouper les travaux scientifiques et de les appliquer afin de résoudre des problèmes spécifiques d'usage pratique. Enfin, le développement expérimental qui est fondé sur des connaissances existantes, consiste en la fabrication de nouveaux produits et dispositifs et à l'amélioration de certains procédés. Récemment, les études théoriques et empiriques s'accordent à montrer que la R&D d'une entreprise ou d'une région peut avoir des effets positifs sur la productivité d'une autre entreprise ou région. Certains auteurs montrent même que les connaissances découlant de la recherche fondamentale, de la recherche appliquée et du développement technologique (Jaffe, 1989), autrement dit, toute activité de R&D peut être source d'externalités (Frascati, 2002).

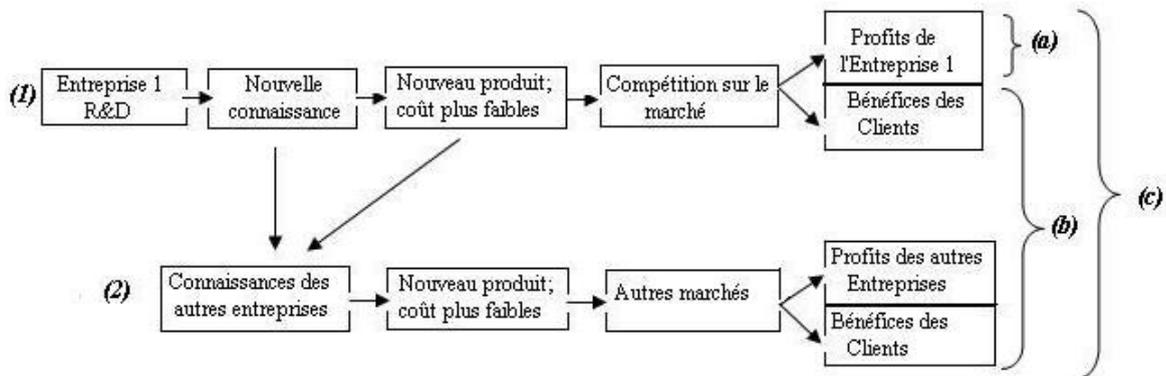
Plusieurs études (Arora et Gambardella, 1994 ; Fritsch et Lukas, 2001 ; Del Canto et Gonzalez, 1999 ; Veugelers, 1997 ; Paul et al., 2000 ; Negassi, 2004) s'attachent à montrer le double rôle de la R&D. D'une part, la R&D permet de générer des innovations. D'autre part, elle permet de développer la capacité de l'entreprise à identifier, assimiler et exploiter les connaissances extérieures (Cohen et Levinthal, 1990 ; Kriaa et Karray, 2010). Ainsi, les entreprises investissent dans leur propre R&D (R&D interne) pour être en mesure d'utiliser les connaissances technologiques disponibles à l'extérieur (R&D externe).

De manière générale, ces études mettent davantage l'accent sur la dimension positive de ces effets de débordement. Ainsi, une innovation dans les inputs permet des réductions de coûts et une innovation dans les produits par exemple favorise l'expansion

8. y compris la connaissance de l'homme, de la culture et de la société.

des marchés. Nous parlons donc d'effets d'externalités de R&D ou d'effets d'externalités positifs.

Nous présentons ci-dessous un schéma établi par Jaffe (1998) qui explique comment les activités de R&D présentent une source d'externalités :



Avec (a): Retombées privées; (b): Gap d'externalités et (c): Retombées sociales

Figure 1.1 – Les retombées privées et sociales : cas des externalités de connaissance

Source : Jaffe (1998)

D'après cette figure, lorsque l'entreprise 1 investit dans la R&D, elle produira une nouvelle connaissance qui peut conduire à une amélioration de la qualité d'un produit et/ou à une baisse de son coût de production et/ou de commercialisation (1). Selon Jaffe (1998), les externalités de connaissance qui vont vers les entreprises agissent aussi sur d'autres marchés que celui de l'entreprise 1 (amélioration de la qualité d'un produit et/ou à une baisse de son coût de production et/ou de commercialisation des autres entreprises) (2). Il y a alors gain pour l'entreprise 1, des gains pour les clients, des gains pour les autres entreprises et d'autres gains pour les clients. Ces externalités représentent donc la différence entre les retombées sociales et les retombées privées. Nous pouvons illustrer l'émergence de ces externalités par l'exemple suivant : une entreprise A entreprend une activité de R&D qui lui permet de créer un nouveau produit. Ce produit peut accélérer l'invention d'un autre produit concurrent dans la mesure où l'entreprise B (l'entreprise concurrente) peut enrichir ses connaissances en étudiant le produit de l'entreprise A : il s'agit donc d'externalités de connaissance profitant à la seconde entreprise B.

1.2.2. La nature tacite de la connaissance comme justification des Localized Knowledge Spillovers (LKS)

Depuis la distinction entre le caractère tacite et codifiée de la connaissance (Polanyi, 1966), les travaux dans le domaine de l'économie de la connaissance introduisent des arguments théoriques montrant que les risques de fuite de cette dernière sont moins forts que prévu et que la connaissance reste souvent appropriable par son créateur pendant de longues périodes (Foray, 2000 ; Cowan et al., 2000).

En effet, la distinction entre connaissance tacite et connaissance codifiée a constitué le second enseignement de l'économie fondée sur la connaissance . Par définition, la connaissance codifiée est parfaitement explicite et fait référence à ce qui est su et qui peut être expliqué. Elle peut être facilement reproduite et stockée à faible coût. Cette connaissance est donc très proche de la notion de l'information (Arrow, 1962).

Contrairement aux connaissances codifiées, les connaissances tacites sont composées de compétences innées ou acquises, de savoir-faire et d'expérience. La dimension tacite implique que la reproduction de la connaissance est une opération coûteuse, voir dans certains cas impossible à réaliser : "*Nous savons plus que ce que nous pouvons dire*" (Polanyi, 1966). La manière la plus simple de définir la connaissance tacite est donc d'indiquer ce qu'elle n'est pas : les travailleurs expérimentés peuvent transférer des connaissances tacites aux novices grâce à des contacts directs et par le dialogue, par exemple, par un encadrement ou par un tutorat. L'exécution réussie d'une compétence dépend donc de "*the observance of a set of rules which are not known as such to the person following them*" (Polanyi, 1958 ; p.49). C'est-à-dire la connaissance qui est "*imperfectly accessible to conscious thought*". (Nelson et Winter, 1982, p. 79).

Foray (2000) précise que les connaissances tacites sont difficilement exprimables hors de l'action du détenteur dans la mesure où le savoir tacite peut même être ignoré par celui-ci. Ainsi, la dimension tacite de la connaissance permet à celui qui la détient d'exercer un certain contrôle puisque seule la démonstration volontaire et l'apprentissage en permettant l'acquisition. Ces types de connaissances ont donc une excluabilité naturelle. Par conséquent, l'échange, le transfert et l'apprentissage d'une connaissance tacite exigent l'implication délibérée et volontaire de celui qui la possède, ce qui explique que les opérations de gestion de la connaissance comme par exemple la mémorisation, la recombinaison et la réutilisation sont complexes et coûteuses.

Depuis les travaux de Marshall (1920) jusqu'aux travaux de la géographie de l'innovation, la dimension tacite des connaissances a largement expliqué le caractère localisé des externalités de connaissance, d'où la notion Localized knowledge spillovers (LKS) ou les externalités de connaissance localisées. Elles ont été définies par Breschi

et Lissoni (2001, p975) comme suit : "*knowledge externalities bounded in space that provide an advantage for closely located firms in accessing critical knowledge sources needed to create innovations at a faster rate than competitors located in other areas*".

De nombreux travaux se sont appuyés sur les LKS pour expliquer la concentration des activités d'innovation et surtout ont constitué un argument éclairant la relation étroite entre espace et innovation (Arrow, 1962; Feldman (1994 et 1996); Caniels, 2000). Dans cette perspective, ces travaux partent de l'idée que si les *spillovers* sont localisés, alors certaines régions doivent bénéficier d'effets cumulatifs favorisant la production d'innovation (Autant-Bernard et al., 2008).

Avec la distinction entre le caractère tacite et codifiée de la connaissance, nous pouvons tirer deux conclusions de ces travaux : la distinction information/connaissance indique que l'étendue des externalités de connaissance est moins importante que celle prévue dans l'approche d'Arrow (1962) puisque la connaissance est une capacité d'apprentissage et de cognition et non pas un flux de connaissance disponible à tous les acteurs. Ensuite, le recours à l'examen des propriétés de la connaissance et la distinction entre sa nature tacite/codifiée montre que les externalités de connaissance codifiées sont plus importantes que celles des connaissances tacites puisque la diffusion de ces dernières est limitée géographiquement. Les deux conclusions convergent vers le même constat : les externalités de connaissance ne sont pas automatiquement générées et sont moins importantes que prévu.

1.2.3. Remise en cause de la dimension tacite des connaissances comme justification des LKS

Comme il a été souligné ci-dessus, l'hypothèse de la dimension locale des externalités de connaissance a pour origine la supposition qu'une partie de la connaissance est tacite et que la proximité géographique (le face à face) est une condition nécessaire à son échange ou à sa transmission. La nature tacite de la connaissance comme justification des LKS a néanmoins été remise en cause par certains théoriciens.

1.2.3.1. Le débat sur la notion de LKS

Perrat (2000) souligne que la focalisation excessive sur la notion de territoire ou d'espace masque les arrangements organisationnels par lesquels les LKS sont gouvernées ainsi que le type des relations qu'entretiennent les acteurs situés au sein même de ces espaces.

Dans la même direction, Breschi et Lissoni (2001) constatent un usage excessif des notions de "*localized knowledge spillovers*" dans les différents travaux de la

nouvelle économie géographique (NEG). Ceci aboutit à une confusion conceptuelle et à l'apparition de certaines incohérences dans les mesures de ces phénomènes supposés généralement liés aux connaissances tacites. Ces auteurs affirment aujourd'hui qu'il est temps de cesser de considérer les externalités locales de connaissance comme une "boîte noire".

Selon Breschi et Lissoni (2001), dès qu'on essaye d'ouvrir "*the black-box of LKSs*", il devient clair que d'un côté, ce qui peut apparaître comme externalités de connaissance, qui échappent à la volonté des acteurs, peuvent être formalisées via des stratégies délibérées, dans le cadre des relations inter-firmes ou entre les universités et les firmes. D'un autre côté, ces externalités, supposées dans la littérature comme étant de nature technologique "pure", peuvent être également pécuniaires et passer par des mécanismes de marché tels que le marché du travail, le marché des technologies et les accords au sein des réseaux.

Gallié (2000) a remis en cause l'existence même des *spillovers*. Pour cet auteur, le phénomène des externalités de connaissance pourrait se réduire et même ne peut plus en découler si nous considérons la distinction faite par l'économie de la connaissance entre connaissance tacite et connaissance codifiée. Selon cet auteur, le problème de diffusion des connaissances codifiées peut être atténué grâce à la mise en place du système de brevets dans la mesure où ces derniers apparaissent comme un instrument incitatif permettant l'appropriation de la connaissance et la protection de l'innovation. Ainsi, la connaissance codifiée et brevetée est appropriable par son inventeur contrairement à la connaissance codifiée non brevetée. Les connaissances tacites, quant à elles, sont personnelles à l'agent. Ces connaissances se prêtent mal à l'échange et à la diffusion car elles sont plus facilement contrôlables (Foray, 2000). Par conséquent, l'adoption des deux définitions strictes des connaissances tacites et des connaissances codifiées ne peut pas émaner des externalités de connaissance : la connaissance tacite requière des interactions sociales pour se transférer et la diffusion des connaissances codifiées est limitée grâce au système de brevet.

En revanche, l'idée que la diffusion des connaissances codifiées est limitée par le système de brevet ne peut pas être validée dans la mesure où l'utilisation de ce dernier présente une certaine limite. Le brevet ne peut pas être considéré comme un moyen efficace d'appropriation des connaissances, ce qui explique le risque des fuites de ces dernières⁹. Dans la même lignée, d'autres critiques ont été avancées en rapport avec la dimension tacite de la connaissance. Ces limites sont présentées dans la sous-section suivante.

9. Voir section suivante pour plus de détail sur le système de brevet.

1.2.3.2. Transmission des connaissances tacites et contact face à face

Les années 1990 ont vu se développer de nombreux travaux empiriques qui se sont concentrés sur la question de l'échelle spatiale des externalités de connaissance (Carrincazeaux, 2009). Ces externalités ont été considérées essentiellement liées au contexte local et l'espace apparaît comme support essentiel de coordination des activités d'innovation des firmes et des régions. Autrement dit, la connaissance tacite est liée à son détenteur et de ce fait, sa transmission ne peut s'effectuer qu'au sein de la firme ou du laboratoire qui a réalisé la découverte. La connaissance tacite a constitué donc un argument sous-jacent qui a longtemps justifié la localisation à proximité pour bénéficier des externalités de connaissance (Breschi et Lissoni, 2001).

Par ailleurs, l'acquisition de la connaissance tacite requiert l'expérience née des relations sociales et d'apprentissage (Foray et Lundvall, 1996). La démonstration par un "*expert*" en la matière, même s'il est situé dans le même lieu que l'apprenti, doit être volontaire. Être géographiquement proche, ne signifie pas nécessairement être en contact. En effet, deux firmes concurrentes par exemple peuvent s'ignorer et se détester ce qui est en défaveur des transmissions des connaissances. Ainsi, une condition nécessaire pour la transmission des connaissances tacites est d'avoir la volonté de la transmettre. La proximité géographique n'apparaît donc pas suffisante pour transmettre les connaissances tacites s'il n'existe pas de canaux de transmission ou interactions entre individus.

Par ailleurs, les externalités de connaissance représentent par définition les connaissances transmises d'une manière volontaire ou non (voir §1.1). La transmission des connaissances tacites nécessite par contre une volonté de la part du transmetteur ce qui ne peut avoir lieu qu'à travers les canaux des transmissions¹⁰. A ce niveau, nous pouvons dire qu'il s'agit bien d'une diffusion des connaissances et non pas d'externalités dans la mesure où avoir la volonté de transmettre l'information est une condition nécessaire pour bénéficier des connaissances tacites.

Mais en revenant à la définition de la connaissance, nous pouvons dire que les externalités de connaissance (si nous supposons que la connaissance ne produit que les externalités positives) englobent le transfert volontaire des connaissances et le transfert non volontaire. Ceci explique peut-être pourquoi la majorité des études utilise la notion externalités même si leur transfert involontaire est trop faible pour ne pas dire inexistant.

10. Voir section 4 : Frontières et canaux de diffusion des connaissances.

1.2.3.3. La codification des connaissances tacites peut atténuer les LKS

Les connaissances ont une dimension tacite. De ce fait, elles posent plusieurs difficultés. La première difficulté concerne l'échange, la diffusion et l'apprentissage de connaissances tacites qui supposent la mobilité et la démonstration volontaire des personnes qui les détiennent ; ces connaissances tacites sont donc coûteuses et difficiles à mettre en oeuvre. De plus, le stockage et la mémorisation des connaissances tacites sont conditionnés par le renouvellement -de génération en génération- des personnes qui les détiennent. Il suffirait d'une coupure d'une génération pour perdre un grand nombre de procédures et d'oublier donc "*comment on fait*" (Mackenzie et Spinardi, 1995). Ceci est appelé aussi "*risque de désinvention*" (Foray, 2000). Une troisième opération difficile à réaliser pour le cas des connaissances tacites, voir même impossible, c'est de classifier ou répertorier systématiquement ces connaissances comme pour le cas des connaissances codifiées. Ces difficultés et ces risques font de la codification de la connaissance un enjeu majeur de l'innovation et du développement économique. La codification est ce qui permet d'exploiter "*cette propriété unique de l'homme, qui est de pouvoir placer sa mémoire en dehors de lui-même*" (Leroi-Gourhan, cité par Favereau, 1998). La codification consiste donc à placer la connaissance sur un support ; elle sera libérée de son rattachement à une personne. La codification permet ainsi, à partir de sa facilité de reproduction, de multiplier les "*copies*" ce qui améliore les conditions de recherche et de transportabilité (Simon, 1982). La connaissance devient donc transférable, indépendamment du transfert d'autres ressources, telles que les personnes qui avaient incorporé les connaissances tacites.

Ainsi, les firmes globales ou multinationales ont recours à la codification pour harmoniser les savoirs et diffuser géographiquement des pratiques nées localement (Foray, 2000). Par conséquent, il n'est pas toujours nécessaire d'être proche géographiquement pour transmettre les connaissances tacites. La connaissance tacite peut être codifiée et transmise par la suite via des canaux de transmission comme par exemple les réseaux sociaux. Ceci remet en cause aussi la dimension spatiale des externalités de connaissance qui a longtemps expliqué les agglomérations des activités innovantes. Donc contrairement à ce qui a été mis en exergue par les géographes, la proximité géographique n'est pas toujours nécessaire pour transmettre les connaissances.

Ainsi, le problème le plus compliqué rencontré dans les études de localisation est la difficulté de séparer les retombées des autres causes qui pourraient expliquer la concentration géographique des activités technologiques. Ce type de savoir, les savoirs tacites, ainsi que les compétences, sont incorporés dans les individus et les organisations et sont, de ce fait, difficilement transférables. Cowan et al. (2000)

rajoutent, en soulignant la nature complexe de la connaissance, l'impossibilité de séparer empiriquement les deux connaissances tacites/codifiées.

1.2.4. La nature variée de la connaissance comme justification des Distant Knowledge Spillovers (DKS)

La notion de "*distant interactions*" a été prise en compte avec les travaux d'Owen-Smith et Powell (2003). Ces auteurs utilisent le terme "*pipeline*" en référence aux canaux de transmissions désignant les interactions distantes.

Dans une étude sur le domaine de la biotechnologie dans la ville de Boston, les auteurs concluent que les externalités de connaissance sont plus importantes au sein d'un réseau que sur l'ensemble de ses frontières. Selon ces auteurs, les entreprises de biotechnologie de Boston ne sont pas seulement intégrées dans des réseaux régionaux d'innovation, mais aussi dans les réseaux sociaux qui ne sont pas bornés géographiquement. Dès lors, l'espace n'est pas le seul déterminant de la diffusion des externalités "*Decisive, non-incremental knowledge flows are often generated through 'network pipelines', rather than through undirected, spontaneous 'local broadcasting'.*" (Bathelt et al., 2004).

Plus tard, Bathelt, Malmberg et Maskell distinguent dans leur article de 2004 les "*pipelines*" des "*buzz*". Contrairement aux pipelines qui se réfèrent à la distance, les "*buzz*" renvoie à "*l'écologie*" de l'information et de la communication créées par les contacts face à face, par la co-présence et la co-localisation des gens et des firmes dans la même industrie ou région : "*This buzz consists of specific information and continuous updates of this information, intended and unanticipated learning processes in organised and accidental meetings*" (Bathelt et al., 2004).

Les "*pipelines*" et les "*buzz*" se renforcent mutuellement. En effet, grâce à leur potentiel dans l'intensification des interactions locales, les pipelines soutiennent la cohésion des clusters et renforcent les relations d'interactions entre acteurs internes (Murdoch, 1995). Ceci constitue un phénomène propice à la création de connaissances. Malmberg et Maskell (2006) s'inspirent ainsi des travaux de Richardson (1972), et expliquent que la création de connaissances prend place verticalement lorsqu'il s'agit des connaissances complémentaires. Quand les acteurs sont similaires, la création de connaissances se fait horizontalement car les acteurs entretiennent des relations de concurrence et celles-ci sont donc produites dans le processus de comparaison, de sélection et d'imitation constante (Malmberg, Maskell, 2005).

Pour aller au delà du dogme selon lequel il est important d'être proche pour innover, il est nécessaire de franchir la simple distinction "*tacite/codifiée*" de la connaissance.

Cowan et al. (2000) distinguent trois types de connaissances selon le degré de visibilité de leur code : la connaissance articulée, non-articulée et non-articulable. D'après ces auteurs, cette dernière forme est marginale, nous nous intéressons donc qu'aux deux premières.

1.2.4.1. Distinction entre connaissances articulées et connaissances non articulées

La connaissance articulée est par définition la partie de la connaissance tacite qui peut être codifiée, et peut donc exister indépendamment de son détenteur sous condition d'investissement dans sa codification. Le code utilisé est parfaitement visible et spécifique à un contexte social. Ce peut être, par exemple un théorème, une circulaire, etc. La connaissance non-articulée ne fait pas référence explicitement à un code. Les auteurs distinguent cette dernière en deux catégories : connaissance non-articulée où le code existe mais est "*déplacé*" et connaissance non-articulée où le code n'existe pas. Pour le premier cas, le code est internalisé seulement par le groupe qui l'utilise. En d'autres termes, des connaissances sont utilisées au cours de conversations sans qu'elles soient définies puisque leur signification est évidente pour tous les agents concernés. Cette connaissance peut donc sembler tacite pour un observateur extérieur (Cowan et al. 2000). Pour le second cas, le code n'existe pas. Cowan et al. (2000) ont établi une autre distinction à ce niveau. Le code n'existe pas et donc la connaissance est tacite mais manifeste dans le sens où elle commence à être exprimée ; lorsqu'elle commence à être divulguée par son inventeur, elle devient une connaissance dont le code n'est pas encore complètement stabilisé. Ensuite, elle peut être codifiée. A terme, son code peut devenir latent.

Toutefois, certaines connaissances peuvent rester tacites, ce qui constitue le deuxième type des connaissances où le code n'existe pas. Cette connaissance "*classique*" est tacite ou latente puisqu'elle ne peut être articulée même avec la volonté du détenteur (Mangolte, 1997). Elle renvoie donc à la partie de la connaissance tacite qui fait référence à des "*routines*" dépendantes de l'action de celui qui la détient, son existence est même ignorée ou négligée (Foray, 2000).

Le tableau (1.1) présente une classification en fonction du degré de codification des connaissances et du degré de visibilité des codes (manifeste ou latent) :

	Connaissance manifeste	Connaissance latente
Connaissance codifiée	Code explicite	Code déplacé
Connaissance non-codifiée	Code non-stabilisé	Pas de code

Tableau 1.1 – Classification des connaissances en fonction de leur degré de codification et de manifestation

Source : Cowan et al. (2000), p 235

Il ressort de cette distinction qu'il n'est pas évident de codifier toutes les connaissances tacites puisque une part de ces connaissances reste toujours non codifiable. La célèbre phrase de Polanyi (1966) éclaire parfaitement cette idée "*we can know more than we can tell*" (M. Polanyi, 1966, p.4). Les individus ont toujours la capacité de savoir plus ce qu'ils peuvent expliquer. L'agent n'a pas toujours conscience des connaissances tacites qu'il possède ce qui n'est pas le cas pour les connaissances codifiées.

Une autre raison qui justifie la difficulté de codifier les connaissances tacites, est le coût élevé que nécessite cette codification. Il est vrai que cette codification permet aux agents d'effectuer un certain nombre d'opérations à un coût marginal très faible. Le transport et le transfert, ainsi que la reproduction des connaissances représentent des fonctions dont les coûts diminueront sensiblement grâce à la codification. En revanche, le coût fixe que supporte la codification est dans certains cas très élevé. Il n'est pas toujours économiquement efficace de codifier une connaissance selon une analyse coût-avantage.

De plus, la connaissance tacite est nécessaire à l'usage des connaissances codifiées et vice-versa. Il y a toujours une part des connaissances tacites dans les connaissances codifiées ce qui montre que ces deux types de connaissances sont complémentaires. Par conséquent, il est difficile de séparer les usages de ces deux types de connaissances même si la plupart des études théoriques font cette distinction tacite/codifiée.

Nous concluons donc que l'explication de la dimension locale des externalités de connaissance ne peut pas s'appuyer uniquement sur leur dimension tacite, pour les deux raisons suivantes : le caractère tacite de la connaissance dépend de l'accès ou non au code. Les connaissances tacites dont l'accès au code est possible ne sont pas nécessairement des connaissances purement tacites et transmissibles uniquement par contact face à face. En effet, le code permet de faciliter leur transmission.

La deuxième explication réside dans le fait qu'aucune connaissance n'est parfaitement appropriable par son inventeur en raison notamment de l'incertitude sur l'usage des connaissances (Gallié et Guichard, 2002). Ainsi, en fonction de l'accès ou

non au code, la connaissance peut apparaître tacite et donc inaccessible à certains, et codifiées à ceux qui possèdent le code. Dans ce cadre, il n'est plus nécessaire d'être à proximité pour capter les *spillovers*. La condition de co-localisation est remplacée entre autre par l'accès ou non au code. En effet, les interactions ne lient pas nécessairement deux agents d'un même territoire s'ils n'ont pas accès au code. La co-localisation ne paraît pas alors nécessaire dans des conditions.

Il résulte de cette analyse que les connaissances tacites et codifiées sont complémentaires et que considérer les externalités de connaissance comme localisées en raison du caractère tacite de la connaissance présente des limites même si cela reste l'explication la plus évoquée dans la plupart des travaux.

1.2.4.2. Nature collective et composite de la connaissance

Le débat sur la nature des connaissances a été abordé de différentes manières dans la littérature économique. Si Cowan et al. (2000) ont montré qu'il est difficile de distinguer les connaissances tacites des connaissances codifiées, Antonelli (2006) montre qu'il est difficile de considérer la connaissance comme un bien public au sens strictement arrowsien.

En effet, plusieurs mécanismes peuvent limiter l'accès et l'appropriation de la connaissance. Ces mécanismes peuvent être institutionnels tels que les droits de propriétés intellectuels ou socioculturels comme le langage. En même temps, il est difficile de considérer la connaissance comme un bien privé parfaitement appropriable à cause des phénomènes d'essaimage et de débordement de la connaissance (Crevoisier et Jeannerat, 2009).

Au delà du débat bien public ou privé, Antonelli (2006) avance que la connaissance peut être abordée comme une activité collective nécessitant la capacité d'entrer en interaction avec différents acteurs au sein des contextes économiques et sociaux. Cette manière différente d'aborder la nature de la connaissance permet de souligner que si la dimension tacite des connaissances est importante dans l'innovation, l'interaction entre agents économiques constitue un facteur plus important de leur transmission. En effet, en considérant que les externalités de connaissance sont localisées, l'augmentation de la distance géographique agit négativement sur leurs transmissions. Lorsque la distance devient plus importante, la diffusion des *spillovers* diminue puis leurs effets de débordement disparaissent. En soulignant le caractère collectif de la connaissance, la transmission de cette dernière peut être réalisée à plus grande distance.

Antonelli et Calderini (2008) parlent aussi du caractère composite de la connaissance, c'est-à-dire se combinant d'avantage au sein de projet industriels complexes : "*knowledge is composite, when it is it self the result of the synthesis of many*

different elementary knowledge modules" (Antonelli, 2006). En effet, les dynamiques de connaissances se transforment de manière trans-sectorielle, autour d'entités composites, comme la santé, la communication ou encore le tourisme (Cooke et Delaurentis, 2007). Les acteurs économiques ont désormais un accès facile à des connaissances nombreuses et dispersées dans l'espace. Ces connaissances vont être articulées avec celles qui sont disponibles ailleurs pour en produire d'autres. Cette vision renouvelée de la connaissance a constitué une caractéristique des approches territoriales d'innovation. Au cours des vingt dernières années, ces approches ont joué un rôle important dans l'économie de l'innovation et de la connaissance. Le territoire constitue pour elles une entité socio-spatiale. Au sein de cet espace, les connaissances sont construites et articulées avec d'autres connaissances pour produire l'innovation. D'où l'importance de ce caractère composite de la connaissance dans la production d'innovation.

Après avoir présenté la relation entre les connaissances et l'émergence des externalités de connaissance, nous discutons dans la section suivante de l'importance de ces deux phénomènes dans les activités innovantes.

1.3. L'innovation : une activité concentrée géographiquement, mais dans des frontières mouvantes

L'innovation est considérée comme le lien entre la création des connaissances et le système de production (Cabrer-Borras et Domingo, 2007). Elle constitue le moteur endogène de la dynamique du système économique (Schumpeter, 1942). Plusieurs travaux se sont focalisés systématiquement sur l'étude des facteurs traditionnels intervenant dans la production d'innovation notamment les dépenses en R&D, la taille de l'entreprise, le capital humain, etc. (Schumpeter, 1942 ; Schwartz, 1975 ; Baldwin et Scott, 1987 ; Cohen et Levinthal, 1989).

Cependant, d'autres analyses ont fait leur apparition par la suite, cherchant à étudier l'impact de l'espace sur l'activité innovante à travers l'étude du concept d'externalités de connaissance localisées LKS. Progressivement, l'espace est considéré comme l'un des facteurs les plus déterminants de l'innovation (Escribano et al, 2009). En effet, ces travaux montrent que l'espace n'est pas neutre et qu'il est intéressant d'étudier la concentration des activités innovantes afin d'apporter un éclairage sur les interactions entre espace et innovation. Pour compléter ces études, nous nous focalisons dans cette section sur l'étude des LKS pour montrer leur rôle dans l'innovation tout en tenant compte des externalités de connaissance non localisées ou distantes (DKS) qui jouent aussi un rôle aussi important.

Dès lors, il devient nécessaire de définir l'innovation et de présenter ses caractéristiques avant de voir comment la dynamique d'innovation est influencée par les LKS et aussi les DKS.

1.3.1. Définition et mécanismes de l'innovation

Schumpeter (1934) a été parmi les premiers auteurs à avoir défini l'innovation. Il la considère comme " *simply the doing of new things or the doing of things that are already being done in a new way*" (Schumpeter, 1934). En d'autres termes, l'innovation représente la valorisation économique (marchande ou non) des nouvelles idées. Les avancées technologiques qui en découlent consistent en l'invention de nouveaux produits, l'amélioration de la qualité ou la réduction du coût de production de biens et services existants.

Schumpeter (1934) a identifié cinq types d'innovation : l'introduction d'un nouveau produit ou le changement des qualités d'un produit déjà existant, l'introduction d'une nouvelle méthode de production ou de transport, l'ouverture d'un nouveau marché, le développement de nouvelles sources de matières premières et de facteurs de production et enfin l'évolution dans l'organisation industrielle. Les deux premiers types d'innovation sont les plus étudiés dans la littérature et sont connus respectivement sous les expressions d'innovation de produit et d'innovation de procédé. La direction générale des entreprises (2006) a défini ces deux concepts comme suit : "*L'innovation de produit se caractérise par l'introduction sur le marché d'un produit (bien ou service) nouveau ou nettement modifié au regard de ses caractéristiques fondamentales, ses spécifications techniques, des logiciels incorporés ou de tout autre composant matériel ou immatériel incorporé, ainsi que de l'utilisation prévue ou de la facilité d'usage. L'innovation de procédé se définit par l'introduction dans l'entreprise d'un procédé de production, d'une méthode de fourniture de services ou de livraison de produits, nouveaux ou nettement modifiés. Le résultat doit être significatif en ce qui concerne le niveau de production, la qualité des produits ou les coûts de production et de distribution. L'innovation peut résulter d'un produit ou procédé nouveau pour l'entreprise mais pas pour le marché. C'est le cas lorsqu'une entreprise adopte une technologie déjà mise en oeuvre par un concurrent*".

Par ailleurs, l'innovation peut aussi être analytiquement décomposée en innovation technologique et innovation non technologique. L'innovation technologique englobe l'innovation de procédé et de produit et l'innovation non technologique comprend, quant à elle, l'innovation organisationnelle et de marketing. Nous parlons aussi d'innovation radicale ou de grande ampleur comme par exemple le micro-ordinateur et d'innovation incrémentale ou de petite taille comme la dernière adaptation d'un logiciel préexistant.

En outre, l'innovation présente quelques spécificités par rapport à l'invention. Alors que l'invention correspond à la phase de génération d'une nouvelle idée liée à un nouveau produit ou à un nouveau procédé, l'innovation représente la phase où ces nouvelles idées sont traduites en des produits ou procédés. Depuis les travaux pionniers de Joseph Schumpeter, l'invention constitue une phase importante de l'innovation. Une fois que l'innovation est produite, elle constituera par la suite un moyen de diffusion du nouveau produit ou dispositif à grande échelle. Nous distinguons donc clairement le passage de l'invention à l'innovation. Dans ces conditions, les relations entre invention, diffusion et innovation sont complexes. En effet, chacune influence les autres par de multiples canaux. Guellec (2009) précise que "*les idées nouvelles permettent la commercialisation des produits nouveaux, lesquels à leur tour suscitent de nouvelles idées; les produits nouveaux sont diffusés, et la diffusion encourage en retour la mise au point de produits nouveaux*".

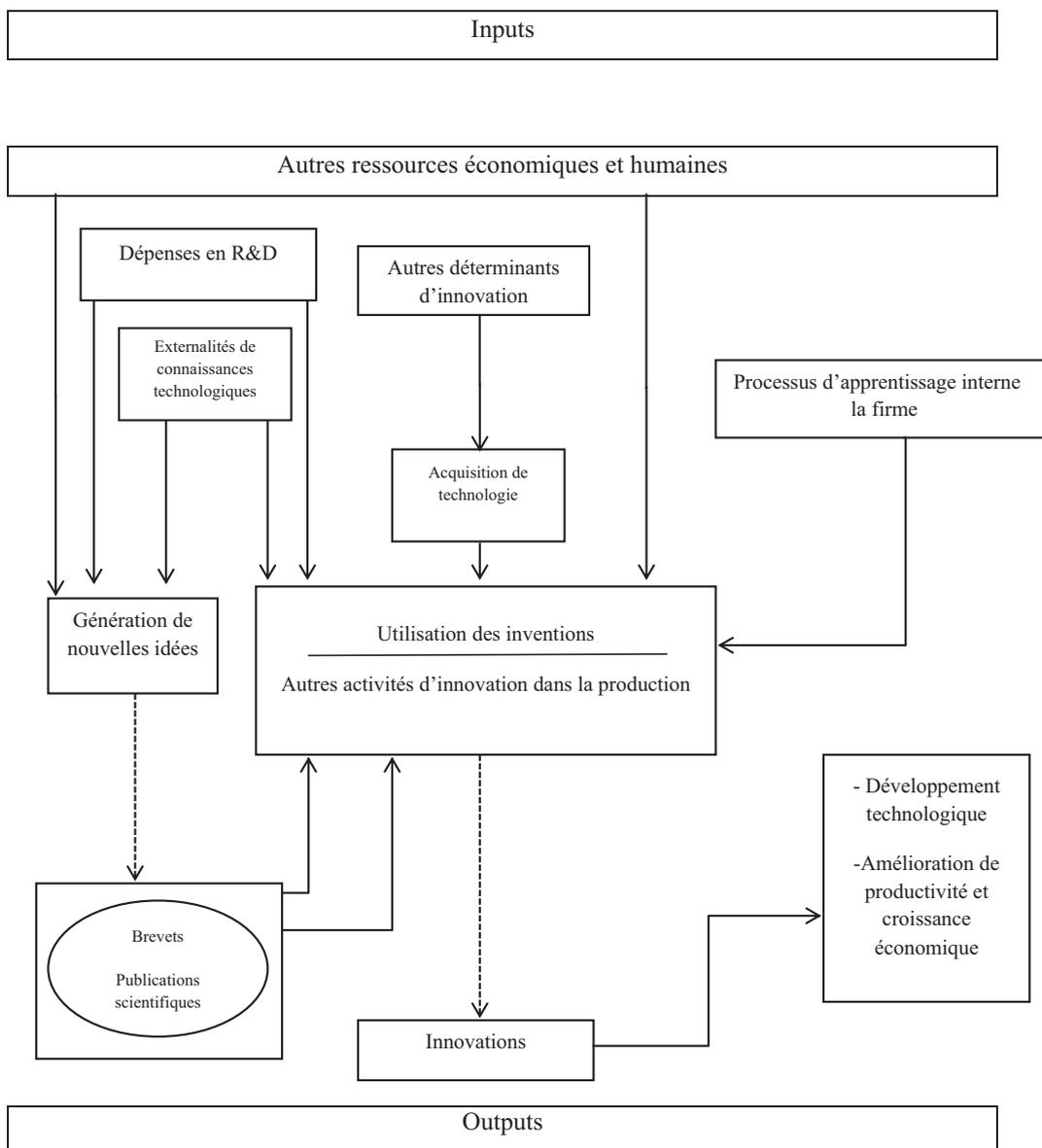
Les activités d'invention et d'innovation sont donc liées même si les inventions ne donnent pas lieu nécessairement à une innovation. En effet, au niveau de chaque étape, le processus d'innovation est l'occasion de multiples boucles de rétroaction. Les bases de connaissances qui représentent la source de l'invention ne sont plus seulement mobilisées lors de la première phase mais tout au long de ce processus.

1.3.1.1. Aspect interactif du processus d'innovation

Dans l'approche évolutionniste (Nelson et Winter, 1982; Dosi et al. 1988), l'explication du phénomène de l'innovation a connu un changement radical. En effet, le changement technologique n'est plus considéré comme exogène. Il est désormais considéré comme un facteur endogène de la croissance. Par ailleurs, chaque agent économique ne jouit que d'une connaissance limitée et possède une perception sélective des choses. Contrairement à l'entrepreneur schumpetérien (lui seul capable de tirer parti des opportunités technologiques qui lui s'offrent), l'échange des connaissances et l'interaction entre ces individus permet de bénéficier des connaissances d'autrui. L'innovation apparaît donc comme le résultat d'interactions de plusieurs acteurs variés. Ainsi, le processus d'innovation est considéré comme "*technology push*" ou comme "*demand pull*". Le premier concept est attribué à Schumpeter (1966) et Rosenberg (1976) et implique que l'innovation est le résultat de la recherche des opportunités technologiques et le deuxième revient à Schmookler (1966) qui précise que l'innovation est fonction de la demande du marché. Les interactions entre acteurs de l'innovation notamment les clients, les fournisseurs, les laboratoires académiques, les acteurs financiers, déterminent l'efficacité du processus d'innovation (von Hippel, 1976).

La figure (1.2) suivante éclaire la relation entre les déterminants de l'innovation et leur rôle dans la dynamique innovante. La R&D est une étape importante en amont du processus d'innovation. Elle représente une source de connaissances. En effet, les scientifiques, ingénieurs, entrepreneurs et inventeurs développent de nouvelles connaissances et s'accaparent les retombées du succès qui découlent de leurs efforts scientifiques. La manière dont les entreprises utilisent la connaissance joue aussi un rôle important dans l'innovation. En effet, bien que les facteurs tangibles tels que les ressources financières et les technologies disponibles soient importantes dans ce processus, la capacité créative dans toutes ses dimensions et le capital intellectuel restent toujours essentiels à l'innovation :

1.3. L'innovation : une activité concentrée géographiquement, mais dans des frontières mouvantes



Source : Borras et al. (2008)

Figure 1.2 – Production et mesure de l'innovation

La figure (1.2) montre que la connaissance présente des caractéristiques particulières dans la mesure où elle peut être considérée à la fois input et output de production de connaissances (Amin et Cohendet, 2004) qui favorisent à leur tour l'innovation. La connaissance est donc un output puisque la production de nouvelles connaissances est le résultat des activités scientifiques et technologiques. Ces nouvelles connaissances peuvent s'ajouter aux autres connaissances déjà existantes pour en produire d'autres, constituant un input de ce processus de production. Le processus de circulation des connaissances pousse ainsi le même agent économique à transformer les connaissances existantes en connaissances nouvelles, et est par conséquent à la fois producteur et consommateur. Ceci crée des rétroactions (feedbacks) dans la mesure où les connaissances produites sont utilisées pour en produire d'autres.

Ainsi, au delà de la séquence traditionnelle du processus d'innovation (la firme fait de la recherche, la recherche mène vers le développement, le développement vers la production et la fabrication de produits qui seront destinés au marché), ces rétroactions jouent un rôle important dans l'innovation. En effet, il est primordial pour l'entreprise d'intégrer ces feedbacks pour pouvoir rapidement améliorer son processus d'innovation notamment dans un environnement de concurrence et d'asymétrie d'information. La production d'innovation peut se révéler sous forme de brevets, de publications par exemple ou bien sous forme de chiffre d'affaires en produits innovants. Ces mêmes indicateurs peuvent être considérés comme moyens de transferts de connaissances. Nous nous limitons dans ce qui suit à présenter l'intérêt des brevets dans le transfert des connaissances. Les critères de choix de cet indicateur, ainsi que la présentation de ses caractéristiques et de ses limites seront présentés dans le chapitre 2.

1.3.1.2. Rôle des brevets dans le transfert des connaissances

Au cours des dernières années, l'importance accordée à l'économie de la connaissance dans le monde entier a conduit à un regain d'intérêt pour les brevets qui représentent un moyen de mesure des activités d'innovations.

Au delà de la fonction de protection de la propriété intellectuelle qui lui est traditionnellement attribuée et en s'inspirant de l'approche évolutionniste de Nelson et Winter (1982), le brevet joue en plus un rôle central de coordination des activités d'innovations. Plusieurs études empiriques et théoriques, réalisées par la suite, insistent à la fois sur son rôle d'incitation et sur son nouveau rôle cognitif et organisationnel.

L'utilisation des brevets comme indicateur de l'innovation remonte aux travaux de Schmookler (1966) et Scherer (1965). À la fin des années 1970, la disponibilité de l'information en forme lisible sur machine auprès de l'office américain des brevets, a permis l'utilisation de ces données avec de plus grands échantillons d'entreprises.

Griliches (1984) suivi par Griliches et al. (1987) et Griliches (1990) ont, par la suite, utilisé ces données dans leur études sur l'innovation. Parallèlement à ces travaux sur les Etats-Unis, d'autres travaux sur l'Europe sont apparus, notamment ceux de Schankerman et Packes (1986) qui ont ouvert la voie de l'utilisation des brevets de l'Office Européen des Brevets (OEB).

En plus de son utilisation en tant qu'un indicateur de l'activité inventive, le brevet est aussi un moyen de transfert des connaissances codifiées¹¹(encadré 1.2). Dans le cadre de la théorie traditionnelle, Arrow (1962) qui a longtemps assimilé la connaissance à l'information, montre que les échanges de connaissances sont bloqués car nul n'accepte de payer pour un bien dont il ignore tout, mais quand on dévoile ce bien à un individu, ce dernier l'acquiert automatiquement sans payer. Pour l'économie de la connaissance, c'est la proximité de la connaissance avec l'information qui empêche son transfert. La nécessité pour les innovateurs est moins de réussir à s'approprier leurs connaissances que de parvenir à les diffuser et à les expliquer aux autres (Callon, 1993). Le brevet facilite donc le transfert des connaissances codifiées qui, ainsi diffusées, permettent la production d'autres connaissances et d'autres idées qui engendrent d'autres innovations.

Encadré 1.2 : Autres fonctions du brevet selon l'économie fondée sur la connaissance

Le brevet joue plusieurs rôles qui justifient sa perpétuelle utilisation au cours du temps (Moussa, 2009) :

1. *Moyen publicitaire et de signalisation des compétences de l'entreprise innovante* : Selon Mazzoleni et Nelson (1998), le brevet joue un rôle de publication dans la mesure où toute innovation brevetée est nouvelle et a une application industrielle. Ceci accroît la réputation de son détenteur et indique à la communauté scientifique et technologique que l'entreprise qui brevète est compétente dans un domaine particulier (Pénin, 2003). Le brevet permet donc à des partenaires potentiels de s'identifier dans un contexte où la multiplicité des acteurs et des informations rend les contacts difficiles (Pénin, 2005).
2. *Instrument pour faciliter le transfert de technologies* : Le brevet peut créer un marché pour l'achat et la vente des technologies. Les entreprises ayant protégé leurs recherches par un brevet vendent les résultats sous la forme de contrat de licence d'exploitation qui spécifie le prix et les modalités de transaction (Cohendet et al., 2006). Ce marché n'existerait pas sans le système de brevet puisqu'il permet, d'un côté, de révéler la

11. Le brevet joue plusieurs fonctions que nous présentons dans l'encadré (1.2). Ces différentes fonctions sont présentées en détail dans *Moussa, I. (2009). Une vision renouvelée du rôle des brevets et des citations de brevets dans le développement, Les Cahiers de l'ATM, n° 24, 51-68.*

description de l'innovation, c'est-à-dire faire de la publicité pour le produit et, d'un autre côté, d'en garantir la protection.

3. *Instrument d'échange de technologies* : Von Hippel (1988) trouve que l'entreprise utilise son portefeuille de brevet afin de mettre en place des accords d'échange mutuel de licences. Dans les domaines où la technologie est complexe, c'est-à-dire où la mise en place d'une innovation nécessite la combinaison de plusieurs technologies, il y a une forte probabilité qu'un individu se retrouve bloqué et n'ait pas accès à une technologie qui lui est nécessaire puisque cette dernière est protégée par un brevet détenu par une autre organisation.
4. *Moyen facilitant la collaboration entre différentes organisations* : La détention d'un brevet donne un signal sur le niveau des compétences. Il permet aux différentes parties souhaitant collaborer de protéger leur savoir antérieur et d'éviter d'être pillées par les partenaires. Cette protection décroît les risques inhérents de la coopération, ce qui naturellement accroît les incitations à participer (Ordover, 1991). Il contribue à définir l'importance de chacun des collaborateurs dans la nouvelle structure, car sans le brevet les participants auraient probablement plus de difficultés à s'accorder sur les modalités de collaboration. Surtout, il facilite la répartition des gains de la collaboration, en permettant le partage des dividendes des partenariats inter-organisations par exemple (Hagedoorn, 2003).
5. *Outil d'interactions entre acteurs hétérogènes* : Le brevet facilite les interactions entre acteurs hétérogènes. Il favorise la rencontre entre son détenteur et d'autres organismes, l'obtention de financement et l'embauche de jeunes chercheurs. Ceci fait de lui un outil central pour rassembler les acteurs autour d'un projet commun.

Pour Cohendet et al. (2006), le rôle du brevet diffère selon la phase de développement de la technologie :

– **Phase d'émergence de la technologie : rôle de coordination**

Lors de cette première phase de développement des innovations, les problèmes d'incertitude peuvent entraîner une diminution massive des échanges entre les différents acteurs de l'innovation. En effet, comme les acteurs hétérogènes ne se connaissent pas, il y a une diminution des échanges. Outre ce risque d'incertitude, l'innovation collective est également freinée par les problèmes de circulation des connaissances dus à l'existence d'une dimension tacite de l'innovation et qui constituent des barrières aux transferts de la connaissance rendant plus compliqué, par exemple, l'accès aux financements nécessaires à la bonne réalisation d'un projet. Dans ce cadre, le brevet constitue un moyen d'assurer les interactions entre les acteurs de l'innovation, il est donc impératif, durant ces premières phases de développement d'une innovation, d'assurer les interactions entre acteurs pour créer des incitations.

– **Phase de développement de la technologie : rôle d'incitation**

Pour Callon (1999), c'est uniquement dans le stade de stabilisation du processus innovatif que le brevet se présente comme un mécanisme d'incitation, dans la mesure où le paradigme technologique atteint un stade où tous les acteurs sont connus, où toutes les opportunités marchandes sont prédictibles, où les langages sont stabilisés, où les expériences de laboratoire sont facilement reproductibles, où la dynamique technologique est bien dessinée à l'intérieur de schémas compréhensibles par le plus grand nombre, où la connaissance est non-rivale et non-exclusive (Cohendet et al., 2006). Le secteur pharmaceutique constitue une rare exception à cette règle (Lieberman et Montgomery, 1988). Ce secteur repose plus sur la science fondamentale, le brevet y est beaucoup plus efficace comme outil d'appropriation des bénéfices de l'innovation.

1.3.2. De l'importance de l'espace pour innover

Jusqu'aux années 1990, plusieurs travaux sur l'innovation ont cherché à étudier les déterminants de la localisation des activités économiques notamment les activités innovantes. Plus précisément, ces travaux ont montré que la localisation des activités de production, la concentration de la main d'oeuvre et la répartition géographique de la demande et des activités de services spécialisés constituent des déterminants de localisation des activités économiques. Pourtant, tous ces facteurs ne suffisaient pas à expliquer cette forte polarisation des activités innovantes qui sont plus concentrées que les activités de production. Sont apparus alors des travaux théoriques et empiriques s'appuyant sur une dimension géographique bornée des externalités de connaissance pour expliquer cette concentration spatiale. Nous pouvons dégager alors une conclusion régulière de ces travaux qui ont essayé d'apporter une explication empirique à la forte polarisation spatiale des activités d'innovation : il paraît évident que l'espace a un rôle à jouer dans la dynamique d'accumulation et de création de connaissances. La proximité géographique influence ainsi le processus d'innovation. Dans un premier temps, les géographes de l'innovation ont souligné la tendance à la concentration spatiale des activités d'innovation (Jaffe, 1989 ; Audretsch et Feldman, 1996). Plus tard, d'autres études sont apparues proposant une explication à cette polarisation spatiale à travers la notion d'externalités de connaissance (Duranton et al., 2008).

Même si Krugman (1991a,b) tenait au début une position différente, les externalités localisées ont longtemps expliqué l'agglomération des activités d'innovation. S'attachant à apprécier l'ampleur de ce qu'Alfred Marshall (1920) avait nommé atmosphère, ces recherches ont confirmé l'hypothèse selon laquelle les externalités de connaissance sont bornées géographiquement (Feldman, 1994a ; Audretsch et Feldman, 1996). En effet, il y avait une liaison étroite entre externalités et espace : l'idée que

l'espace influence les flux de connaissance forme un acquis stable dans les différentes théories en dépit de différences de méthode de mesures et de données utilisées. Les économies d'agglomération représentent des avantages comparés en termes de productivité que procure à une firme ou un ensemble de firmes une région par rapport aux autres, du fait de sa taille et de sa structure (Catin et al., 2007). Les LKS trouvent leurs sources dans l'approche fondatrice de Marshall (1920) qui précise les différentes forces d'agglomération des activités économiques. Les LKS constituent, en plus des rendements d'échelle croissants et de la main d'oeuvre qualifiée, une troisième explication des facteurs d'agglomérations. Ces facteurs ont été identifiés par les travaux de la nouvelle économie géographique (NEG), suite à l'article fondateur de Krugman (1991a,b) qui avait pour objectif d'expliquer les forces d'agglomération des activités économiques en tenant compte de leur inégale distribution spatiale.

Par ailleurs, le modèle de Paul Krugman (1991a) s'inspire de l'approche de Marshall pour insister sur les deux premières forces d'agglomération et négliger la troisième. Les apports de la modélisation par la NEG permettent de distinguer les facteurs d'agglomérations en deux facteurs : facteurs traditionnels qui ne tiennent pas compte des externalités de connaissance et facteur externalités de connaissance dont l'analyse paraît la plus fragile.

1.3.2.1. Le rôle secondaire des LKS dans les premiers modèles d'agglomération de P. Krugman (1991a,b)

Afin de décrire les raisons de la localisation des activités productives dans l'espace, Krugman développe un modèle centre-périphérie sur la structure de Dixit, Stiglitz, Krugman (Krugman, 1979 ; 1980 ; 1981 ; Dixit et Norman 1980 ; Helpman et Krugman, 1985). Ce modèle aborde les questions de localisation, de développement, de régionalisation et d'intégration économique tout en tenant compte des conclusions de la nouvelle théorie du commerce international. Krugman (1991a,b) considère que les externalités de connaissance ont un rôle beaucoup moins important dans l'agglomération que les rendements d'échelle croissants et la main d'oeuvre qualifiée :

- Les rendements d'échelle croissants et le coût de transport représentent un facteur important expliquant la polarisation des activités économiques. Le modèle fondateur de la nouvelle économie géographique proposé par Krugman (1991a,b) apporte une explication à la concentration des activités à travers un arbitrage entre d'une part, le gain entraîné par la concentration de la production afin de bénéficier de coûts marginaux et moyens plus faibles et, d'autre part, les coûts de transport des lieux d'extraction des matières premières jusqu'aux lieux de production et des lieux de production jusqu'aux lieux de consommation.

En effet, lorsque les coûts de transport sont importants, les localisations les plus attractives sont celles qui présentent des rendements d'échelle croissants au niveau de la production. En présence d'une demande importante et face à des coûts de transport relativement faibles, les firmes se laissent guider dans leurs stratégies de localisation par des "*home-market effects*" (Simonis, 2002). Autrement dit, quand les coûts de transport sont faibles, les entreprises peuvent bénéficier des avantages liés à la concentration de la production sans craindre des pertes de parts de marché sur les marchés où se trouve la demande. La demande peut donc jouer un rôle dans l'agglomération des activités innovantes. "*Une entreprise isolée devra payer plus pour se faire livrer des composants électroniques de haute technologie dont elle n'utilise qu'une petite quantité. En revanche, si plusieurs entreprises utilisant ces composants s'installent au même endroit, et induisent ainsi une demande plus importante, le fournisseur va pouvoir livrer ces inputs à un prix plus intéressant, du fait d'économies d'échelle ou parce qu'il va lui-même venir se localiser au même endroit, ce qui diminue les coûts de transport*" (Duranton et al., 2008).

Dans ce contexte, les rendements d'échelle croissants (ou les fortes indivisibilités dans la production) ainsi que la demande locale expliquent la concentration des activités de production.

- La main d'oeuvre qualifiée constitue un deuxième facteur qui explique et favorise le processus d'agglomération des entreprises. En effet, les entreprises cherchent à se localiser aux endroits où il existe une offre constante de main d'oeuvre dont le savoir et les compétences sont spécifiques. Certains auteurs considèrent que les villes sont l'endroit idéal qui permet la formation des individus : "*Cities are places where young people go to learn skills*", Glaeser (1998, p148). Afin de bénéficier de la main d'oeuvre qualifiée, les entreprises ont alors intérêt à se localiser à proximité de ces zones, ce qui permet la constitution d'un vaste marché du travail et le développement de qualifications spécifiques. De la même manière, les travailleurs se dirigent généralement là où ils ont le plus de chance de trouver un travail correspondant au mieux à leurs attentes (Krugman, 1991a,c).

Ainsi, plusieurs études empiriques ont montré que la mobilité des chercheurs et des techniciens au niveau local constitue un vecteur de diffusion des connaissances entre firmes (Saxenian, 1994). La diffusion des connaissances de manière informelle entre les différents chercheurs ou ingénieurs par exemple, constitue une source d'économies de coûts de transaction liées à l'échange d'idées et de connaissances.

Cependant, la mobilité excessive des travailleurs peut impliquer un risque de divulgation des secrets scientifiques et technologiques des firmes (Fosfuri et

Ronde, 2004 ; Motta et Ronde, 2002). Ainsi, en cas de hausse des salaires due à une demande accrue de main d'oeuvre ou à la concurrence entre entreprises, ces dernières seront soumises à un arbitrage entre les avantages liés à la disponibilité d'une réserve de main d'oeuvre qualifiée et l'inconvénient de subir les effets de débauchage des employés (Combes et Duranton, 2006).

1.3.2.2. La prise en compte de l'importance des LKS comme facteur d'agglomération

Plus récemment, Krugman a reconnu lui-même l'importance des externalités technologiques et considère aujourd'hui que les externalités "pures" du type knowledge spillovers sont déterminantes (Krugman, 1991a,b). Par exemple, dans la théorie du commerce international, il affirme que les échanges produisent des externalités positives parce que les entreprises innovantes ne parviennent pas à conserver le monopole des connaissances qu'elles produisent. En revanche, il donne trois raisons pour lesquelles les externalités technologiques sont exclues de l'analyse.

La première raison montre qu'en dépit du succès des clusters de la Silicon Valley ou celui de la route 128 à Boston, il existe d'autres clusters où les externalités de connaissance ne présentent pas un enjeu principal comme par exemple la finance à New York ou les producteurs de bijoux autour de Providence.

La deuxième raison réside dans le fait que les externalités pécuniaires sont plus en mesure d'expliquer l'agglomération des activités économiques que les externalités de connaissance¹². Ces externalités pécuniaires envoient à des interactions qui transitent par le marché (Scitovsky, 1954) et sont considérées depuis les analyses des districts marshalliens comme décisives (Krugman, 1991b). En effet, dans le cadre d'une concurrence imparfaite et de rendements d'échelle croissants, une firme pourrait affecter la demande d'un produit d'une autre firme dont son prix de vente est supérieur à son coût marginal. Cette situation donne lieu à des externalités pécuniaires qui expliquent à leur tour les structures d'agglomérations des activités économiques. Elles sont introduites par la suite dans les modèles théoriques dont le but est de justifier la localisation des firmes dans certaines régions.

La troisième raison qui explique la position de Krugman est liée à la difficulté à reconnaître les externalités. Pour lui, les flux de connaissance sont invisibles, ne laissent pas de traces et ne peuvent être mesurés : *"knowledge flows, by contrast, are invisible; they leave no paper trail by which they may be measured and tracked, and there is nothing to prevent the theorist from assuming anything about them that she*

12. Voir paragraphe origine des externalités de connaissance pour une définition des externalités pécuniaires.

likes" (Krugman, 1991a : 53-54). Le recours aux externalités technologiques n'est donc pas nécessaire à l'explication des mécanismes d'agglomération et l'essentiel des travaux de la NEG s'est concentré sur les modèles avec mécanismes de rétroactions positives. La minimisation de l'importance de ces externalités amène Krugman à interpréter la concentration moderne comme les forces de localisation présentes au XIXe siècle. Il découle de ce constat une remise en cause de la capacité des modèles de la NEG à prendre en compte les spécificités des agglomérations et de les intégrer dans un modèle centre-périphérie (Coissard, 2007).

Même si la NEG montre que les mécanismes d'agglomération dépendent de la forme spatiale des externalités technologiques, leur analyse reste fragile sur cette dernière (Carrincazeaux, 2011). Cette position de la NEG a été contestée par l'approche de la géographie de l'innovation, en particulier Jaffe (1989) pour qui la connaissance laisse des traces par les brevets (Jaffe et al., 1993). Il explique ses réserves initiales par le manque de théorisation de ces effets (Fujita et Krugman, 2004). Dans l'ensemble, ces travaux s'accordaient sur l'importance de la proximité géographique qui paraît décisive dans les activités économiques. Par le biais d'une proximité géographique, les agents économiques vont bénéficier des *spillovers* qui sont à l'origine de ces effets d'agglomération, permettant ainsi d'expliquer pourquoi les activités économiques se localisent dans certains lieux et pas dans d'autres et pourquoi la distribution spatiale des activités économiques est inégale.

Toutefois, en se référant en permanence aux externalités de connaissance localisées, les travaux empiriques et théoriques expliquant l'importance de l'espace dans l'innovation ont tendance à surévaluer ce phénomène.

1.3.3. Une discussion de la nature et de l'importance des LKS dans la production d'innovation

Les travaux sur les externalités de connaissance localisées, notamment l'approche de la géographie de l'innovation, ont largement souligné le rôle de l'espace dans la production d'innovation. La majorité de ces travaux montrent l'importance d'être à proximité pour bénéficier des LKS. Beaucoup s'accordent, à l'instar de Marshall qui considère que "*les externalités sont dans les airs*", sur le fait que les externalités de connaissance issues de la recherche publique ou privée sont locales. Il suffirait d'être à proximité de la source émettrice de connaissances (que ce soit une firme, une université ou un département,...) pour capter ces LKS.

Toutefois, malgré la diversité des travaux proposés en faveur de cette thèse, il semble réducteur d'attribuer l'effet positif de l'espace à la seule existence des *spillovers* de connaissances.

1.3.3.1. L'espace n'est pas structuré seulement par les LKS

Dans la même lignée que les travaux sur l'agglomération des activités innovantes, Tappeiner et al (2008), sur un échantillon de 51 régions européennes, ont cherché à expliquer si l'effet positif de l'espace sur l'innovation est causé seulement par les LKS. Ces auteurs ont exploré cette question en testant les deux hypothèses suivantes :

- *Première hypothèse* : l'autocorrélation spatiale¹³ d'une région quelconque est la résultante des LKS en provenance des régions voisines fortement innovantes.
- *La deuxième hypothèse* : l'autocorrélation spatiale est le résultat de la concentration dans l'espace des facteurs de production. Ces facteurs sont divisés en deux catégories. La première catégorie regroupe les facteurs traditionnels tels que la R&D, le capital humain, etc. et la deuxième catégorie concerne le capital social. Ce facteur d'innovation a été développé par Hauser et al. (2007). En effet, ces derniers ont intégré dans leur modèle une variable proxy mesurant le capital social qui consiste à tester l'existence et l'impact des canaux de transmission des connaissances tacites. Ils mesurent cette variable en utilisant 5 facteurs comme indicateurs du capital social : l'intérêt politique, liens d'amitiés, activités associatives, le degré de confiance et le développement technologique et autonome. Les résultats trouvés montrent que l'autocorrélation spatiale disparaît lorsque le capital social est intégré dans le modèle.

Il découle de cette étude que les LKS ne jouent pas en soi. Son rôle s'atténue au fur et à mesure de l'introduction des variables sociales dans le modèle comme par exemple les relations interpersonnelles. Ainsi, Les LKS ne représenteraient pas la seule explication de l'effet positif de l'espace sur l'innovation, les relations sociales qui se développent dans un milieu géographique sont parfois plus importantes que les LKS.

1.3.3.2. L'innovation est liée aux LKS mais aussi aux DKS et à des proximités non géographiques

Nous avons précisé précédemment que l'effet positif de l'espace sur l'innovation ne s'explique pas seulement par les LKS. Cet effet serait conditionné aussi par l'existence d'interactions effectives entre agents. Toutefois, si les activités innovantes sont concentrées c'est que nous pouvons accepter l'hypothèse d'un bornage géographique des externalités de connaissance. Les externalités de connaissance ne se diffusent pas donc uniformément dans l'espace et le caractère localisé de ces externalités constitue bien une explication, parmi d'autres présentées précédemment, de la forte concentration de l'innovation.

13. Voir chapitre 3 et 4 pour plus de détail sur la notion d'autocorrélation spatiale.

En plus des LKS, d'autres externalités qui ne sont pas nécessairement localisées ont un effet sur la production d'innovation. Ces externalités de connaissance non localisées, appelées aussi externalités de connaissance distantes ou Distant Knowledge spillovers, jouent un rôle important dans la production d'innovation. Ces dernières ne sont pas nécessairement liées à l'espace puisqu'elles se transmettent entre entreprises ayant une certaine proximité non géographique.

Suivant la typologie de Boschma (2005), quatre formes de proximités non géographiques sont distinguées¹⁴ : la proximité cognitive, la proximité organisationnelle, la proximité sociale et la proximité institutionnelle (Balland, 2010). Ces proximités non géographiques favorisent amplement la diffusion des connaissances. En effet, les chercheurs d'une même discipline par exemple, développent le même langage et les mêmes codes indépendamment de leurs positions géographiques. Ainsi, l'importance de la proximité géographique est relativisée puisque les externalités de connaissance se diffusent indépendamment de la distance qui sépare les agents économiques produisant les externalités de connaissance (Boschma et Frenken, 2010 ; Balland et al., 2011). Ces dernières ne sont pas non plus bornées géographiquement et peuvent se diffuser à une distance plus longue que les LKS. Par ailleurs, elles constituent, elles aussi, une explication de la forte agglomération des activités innovantes.

Par ailleurs, nous pouvons distinguer les flux de connaissance (DKS ou LKS) en deux catégories : externalités MAR et externalités Jacobs.

Les externalités MAR : Glaeser et al, (1992) en référence aux auteurs A. Marshall (1980), K. Arrow (1962) et P. Romer (1986), ont élaboré la conception "*MAR*". Ces externalités marshalliennes s'inscrivent dans la logique des économies de localisation intra-industrielles. Le modèle Marshall-Arrow-Romer formalise le raisonnement selon lequel la concentration des industries dans une ville favorise les externalités de connaissance entre firmes et, par conséquent facilite l'innovation dans la "*ville-industrie*" (Maftai, 2010). Selon l'hypothèse soutenue par ces auteurs, les externalités de connaissance ne sont limitées qu'à leur secteur, ainsi leur diffusion vers d'autres industries est supposée inexistante ou, du moins, insignifiante (Feldman et Audretsch, 1999). Les externalités MAR sont donc des externalités intra sectorielles puisqu'elles touchent les entreprises appartenant à un même secteur et augmentant avec l'importance des secteurs économiques dominants.

En revanche, pour d'autres auteurs, les externalités résultent d'un regroupement d'activités différentes. C'est le point de vue de Jacobs (1969) qui montre que les

14. Les proximités non géographiques sont définies et présentées en détail dans la section suivante de ce chapitre.

transferts de connaissances les plus fréquents sont interindustriels : ce sont des externalités de type "Jacobs" (1969).

Externalités Jacobs : La diversité des industries concentrées géographiquement stimule l'innovation. Ces économies externes à la firme et externes à l'industrie à laquelle appartient la firme, sont liées à la diversité dans le milieu régional. Ainsi, les externalités Jacobs résultent d'un regroupement en un même lieu de firmes d'industries différentes. L'impact de ces externalités dépend des caractéristiques propres à chaque secteur (Catin et al. 2007). Les externalités Jacobs sont appelées externalités intersectorielles dans la mesure où elles possèdent un caractère transversal et sont liées à l'importance économique et démographique du territoire considéré. La diversité semble ressortir plus nettement comme déterminant important de la croissance dans les secteurs de haute technologie alors que la spécialisation aurait plutôt un effet d'entraînement dans les secteurs traditionnels.

Les externalités intrasectorielles comme intersectorielles diminuent très fortement avec la distance. Néanmoins, la dimension géographique des *spillovers* est généralement plus marquée pour les externalités intersectorielles que pour les externalités intrasectorielles. En effet, les conclusions énoncées par les géographes montrent que la proximité géographique augmente les échanges intersectoriels de connaissances. Partant de ce fait, nous analysons dans la section suivante les canaux de transmission des connaissances.

Toutefois, Frenken et al. (2007) précisent que cette littérature souffre de deux limites. Tout d'abord, nous pouvons nous demander si les externalités de connaissance sont censées avoir lieu entre secteurs qui ne sont pas liés, comme le suggère la littérature sur les externalités de type Jacobs. Par exemple, il est difficile pour un éleveur de porcs d'apprendre à partir d'une entreprise sidérurgique en dépit du fait qu'ils soient voisins (Asheim et al., 2007). La transmission de connaissances se fait donc entre secteurs ou la distance cognitive n'est pas trop large (Nooteboom, 2000). Deuxièmement, une économie diversifiée peut également agir comme un absorbeur d'un secteur spécifique ou d'un choc asymétrique. Par conséquent, la notion d'externalités de Jacobs peut être dans certains cas confondue avec les externalités positives de la diversification des portefeuilles d'activité (Frenken et al., 2007).

Dès lors, il est essentiel de distinguer les différentes formes de diversité régionale puisqu'elles impliquent des effets économiques différents. Les connaissances doivent se répercuter d'un secteur à un autre quand elles sont complémentaires en termes de bases de connaissances et de partage des compétences. En effet, ces externalités sont plus importantes entre industries commercialisant des produits similaires et partageant

la même technologie et les mêmes compétences (Mameli et al., 2012). En tant que tel, ces connaissances nécessitent "*une variété reliée*" ou une certaine variété connexe afin de permettre des connexions efficaces. Asheim et al, (2007) définissent ce concept comme étant les secteurs qui sont liés en termes de partages et de connaissances complémentaires. En d'autres termes, il y a un certain degré de proximité cognitive afin de s'assurer que la communication efficace et l'apprentissage interactif auront lieu mais sans trop de proximité cognitive afin d'éviter ce que Nooteboom (2000) appelle "*verrouillage cognitif*". Ce concept de "related variety" prend une place importante dans les nouvelles approches (Frenken et al., 2007), l'Italie (Boschma et Iammarino, 2009 ; Quatraro 2010), la Grande Bretagne (Bishop et Gripaos, 2010), l'Espagne (Boschma et al., 2011) et l'Allemagne (Brachert et al., 2011). En effet, et comme l'a indiqué Boschma et al.(2012). "*the more variety across related sectors in a region, the more learning opportunities there are for local industries, the more inter-sectoral knowledge spillovers are likely to take place, and the higher the economic performance of regions*".

Pour ces auteurs, ce n'est pas la diversité régionale (peu de proximité cognitive) ni la spécialisation (trop de proximité cognitive) qui stimule les innovations réelles mais c'est plutôt la spécialisation régionale dans la "*variété reliée*" qui est plus susceptible d'induire un apprentissage interactif et des innovations.

Par conséquent, ce concept dépasse la dichotomie traditionnelle des économies de localisation et des externalités de connaissance.

1.4. Frontières et canaux de diffusion des connaissances

Nous avons présenté précédemment les principales caractéristiques de la connaissance, notamment ses caractères cumulatif et composite, qui expliquent le fait que la contrainte de proximité géographique peut être levée. Il importe donc de préciser les différents canaux de transmission des connaissances qui expliquent le desserrement du rôle de l'espace.

1.4.1. Les caractéristiques des territoires comme déterminant des frontières de la diffusion des connaissances

Pour expliquer la localisation des activités de production, Krugman (1991a) a repris la contribution de Marshall, notamment l'approche par les économies externes développée en 1919. Comme nous l'avons précisé précédemment, la nouvelle économie

géographique cherche à décrire le phénomène de localisation industrielle en faisant appel aux concepts de concentration et d'économie d'échelle, à leur nature et à leur localisation. Néanmoins, cette littérature (Krugman, 1991a,b ; Porter 1990 et 1998) se limite à la prise en compte de facteurs essentiellement économiques. D'autres approches ont été développées récemment, basées sur l'étude des spécificités territoriales et socioculturelles pour expliquer le phénomène d'agglomération.

Dans cette perspective, nous nous penchons dans cette partie sur l'étude du rôle du territoire dans le transfert des connaissances et nous tentons de montrer en quoi les frontières des connaissances sont devenues non spatiales au sein d'un territoire donné.

1.4.1.1. Les Systèmes Territoriaux d'Innovation (STI)

Avec l'avènement de l'économie de la connaissance et depuis les travaux de Marshall (1890), les STI connaissent un succès mondial notamment depuis le triomphe de la Silicon Valley et de la Route 128 aux Etats Unis (Saxenian, 1994). Même si la proximité géographique reste un facteur capital à l'émergence de ces systèmes, nous montrons dans ce qui suit que les autres formes de proximités non spatiales jouent un rôle important au sein de ces unités.

Le concept de "*Systèmes Territoriaux d'Innovation*" (STI) constitue un élargissement du concept "*Systèmes Territoriaux de production*" (STP) développé depuis les années 70-80. En effet, la globalisation de l'économie, l'ouverture des marchés et la nouvelle répartition spatiale des activités dans l'espace ont conduit à un phénomène d'agglomération des activités de production d'où le concept de STP (Aydalot, 1984). Le STP est un ensemble d'unités productives (entreprises, centres de recherches et de formation, organisme d'interface, etc.) qui entretiennent des rapports marchands et non marchands (Gilly, 1987) et qui génèrent une dynamique productive d'ensemble (Maillat, 2003).

Depuis, plusieurs approches et définitions se sont développées autour du concept de STP en raison de leur impact positif sur la performance des entreprises, sur la croissance économique régionale et sur la compétitivité des pays (Rocha, 2004). En effet, que nous parlions de districts technologiques (Antonelli, 1994), milieux innovateurs (Ratti et al. 1997), technopoles, parcs scientifiques (Monck et al., 1988), clusters ; un consensus reconnaît à ces "*modèles d'innovations territoriaux*" (Moulaert et Sekia 2003) leur importance dans la mise en évidence du rôle clé des régions dans la dynamique d'innovation (Asheim et al., 2007).

Les districts industriels : les premières études sur les STP se sont fondées sur la notion de district industriel marshallien. A partir de son étude sur les petites entreprises du secteur de la coutellerie en Angleterre, Alfred Marshall (1890) observe

deux schémas d'organisation industrielle de la production. Le premier schéma est caractérisé par un commandement unique et une forte division du travail intégré au sein d'une grande entreprise et le second est caractérisé par une coordination d'une division sociale du travail désintégré entre firmes plus petites se spécialisant dans un segment du processus productif (Chabault, 2011). C'est autour de ce deuxième type d'organisation que Marshall articule la notion de district industriel défini comme "*un système productif géographiquement localisé, caractérisé par un grand nombre de firmes qui sont impliquées aux différents stades de production d'un produit homogène de manière diverses*" (Pyke et Sengenberger, 1990).

Marshall met en avant la notion "*d'atmosphère industrielle*". En effet, quand des agents économiques sont engagés dans des tâches similaires au sein du même territoire, "*la compétence requise dans leur travail est dans l'air et les enfants la respirent en grandissant*" (Whitaker, 1975). Le district est caractérisé donc par une forte interdépendance entre ses différents acteurs.

En effet, chaque entreprise a tendance à se spécialiser dans une seule ou quelques-unes seulement des phases du processus productif spécifique du district. Ce district représente un cas concret de division du travail localisé. C'est ainsi un mélange de concurrence-coopération entre les PME qui le constituent (Courlet, 1999). Le district industriel repose donc sur plusieurs attributs notamment les interactions et les relations sociales. Il est une entité "*socio-culturelle*". Les travaux sur les districts industriels ont été réactualisés avec les chercheurs italiens (Bagnasco, 1977 ; Trigilia, 1986), plus précisément avec les travaux sur le Nord-Est et sur le centre de l'Italie, en insistant sur les caractéristiques sociologiques et culturelles de ces régions comme facteur explicatif de la dynamique de développement (Courlet, 1991).

Toutefois, le district industriel représente une forme ancienne de réseaux. Cette organisation n'est présente que dans certaines industries traditionnelles comme le textile, la chaussure, etc. Les organisations les plus récentes, notamment celles de hautes technologies sont caractérisées par des relations entre les grandes firmes, les PME et les universités.

Le technopole : ou parc technologique est une forme spécifique de district industriel classique. Il est caractérisé par une forte concentration des laboratoires de recherche publics et privés, par l'hétérogénéité des acteurs et leur spécialisation dans des domaines nouveaux de haute technologie. Le technopole est appelé aussi parc scientifique dans la mesure où il permet de rassembler des firmes de haute technologie au sein d'une aire géographique donnée. En outre, il facilite les relations d'échange entre les acteurs notamment les firmes, les laboratoires de recherche et les réseaux de

coopération. Le concept de technopole repose principalement sur l'hypothèse selon laquelle la proximité géographique facilite la communication entre les différents acteurs du district. En France, les technopoles sont la résultante de politiques publiques visant un développement économique axé sur les activités de haute technologie (Longhi et Quéré, 1991).

Un des objectifs premiers des technopoles est la commercialisation des surfaces industrielles (Doloreux, 2002) et non pas la recherche de liens entre les entreprises. En effet, même si les technopoles sont caractérisés par un nombre élevé d'entreprises innovantes, en particulier des PME issues de secteurs high tech, ces dernières sont rarement engagées dans des réseaux d'innovation.

Les clusters : font référence aux vertus du regroupement et de l'interconnexion de différentes firmes au sein d'un même espace géographique (Porter, 2000 ; Feser, 1998). Porter a défini les clusters comme suit : "*Clusters are geographic concentrations of interconnected companies, socialized suppliers, service providers, firms in related industries, and associated institutions (for example, universities, standards agencies, trade associations) in a particular field that compete but also cooperate. Clusters or critical masses of unusually competitive success in particular business areas, are striking feature of virtually every national, regional, state and even metropolitan economy, especially in more advanced nations*" (Porter, 1998, p197).

Pour le cas des clusters aussi, la proximité géographique occupe une place importante dans la mesure où elle permet la circulation de connaissances, la diffusion d'une culture commune ou la constitution d'un capital humain ou social de l'aire considérée. Parmi les clusters les plus connus nous trouvons les clusters d'innovation. La route 128 et la Silicon-Valley (Saxienian, 1994) représentent les cas les plus illustratifs au sein desquels se trouvent co-localisées de nombreuses entreprises high tech et pour lesquelles les auteurs suggèrent que la concentration de la R&D industrielle et la combinaison de connaissances tacites et codifiées, jouent un rôle majeur.

L'intérêt de ces analyses réside dans la présentation des conditions institutionnelles d'une production d'innovation au niveau régional mais elles ne montrent pas trop comment la proximité géographique permet de créer des liens sociaux, la confiance ou les valeurs communes nécessaires à la transmission des connaissances.

L'approche par les milieux innovateurs : cette approche a pour objectif de comprendre comment et pourquoi certaines régions se révèlent plus innovatrices que d'autres (Maillat et al., 1993 ; Ratti et Bramanti 1997 ; Crevoisier et Camagni 2000 ; Crevoisier 2001). Les milieux innovateurs représentent un ensemble d'acteurs localisés dans un même territoire dans lequel les interactions se développent de

manière multilatérale et sont génératrices d'externalités spécifiques à l'innovation. Ces approches en termes de milieu innovateur mettent l'accent sur le rôle de la dimension territoriale dans l'analyse de la dynamique économique et des processus d'innovation (Maillat, 1995). Plus précisément, une telle approche met en avant les mécanismes par lesquels la proximité spatiale des différents partenaires permet l'émergence, le maintien et le renouvellement d'une dynamique d'apprentissage (Gallié, 2004).

A une nuance près, tous ces concepts présentent le même phénomène. En effet, il s'agit d'une concentration d'agents économiques (concurrents et complémentaires) sur un territoire géographiquement limité, avec des interactions entre elles. Ces agents sont spécialisés dans un secteur d'activité, un service ou une technologie particulière. Ils peuvent s'appuyer sur une structure d'animation associant les acteurs du territoire.

Le pôle de compétitivité : le pôle de compétitivité est apparu plus récemment en France. Cette politique publique est fondée sur le même principe de regroupement d'acteurs économiques (entreprises, centres de recherche, sociétés de services, structures publiques de valorisation, etc.) localisés sur un même territoire, mais sans être nécessairement des regroupements sectoriels. Nous les définissons comme suit : un pôle est par définition *"une combinaison géographique donnée d'entreprises, de centre de formation et d'unités publiques ou privées engagés dans une synergie autour de projets communs innovants. Ce partenariat s'organise autour d'un marché et d'un domaine scientifique qui lui est attaché, et doit rechercher une masse critique pour une compétitivité et une visibilité internationale"* (Source : Site interne de la Diact/ex Datar). Ces pôles semblent également occuper une position prééminente, dans la mesure où ils constituent les principales sources des nouvelles idées et connaissances qui favorisent le développement du processus d'innovation.

En France, l'idée de pôles de compétitivité s'est développée dans les années 1970 sur les modèles de technopoles, à l'instar du Japon, et dans une certaine mesure des États-Unis, telle que la Silicon Valley. Ces pôles favorisent la croissance de la capacité d'innovation, la compétitivité de l'économie française et l'attractivité du territoire national. Ils favorisent également l'emploi et la lutte contre les délocalisations. Le label "pôle de compétitivité" est attribué par une décision du comité interministériel d'aménagement et de compétitivité des territoires (CIACT). Le CIADT du 12 juillet 2005 avait labellisé 67 pôles de compétitivité sur les 105 dossiers de candidature reçus à cette date. Après la réception de nouvelles candidatures et la fusion de pôles déjà labellisés, les pôles de compétitivité sont 71 depuis le CIACT du 5 juillet 2007. Ces pôles sont répartis en trois catégories : des pôles mondiaux, à vocation mondiale et d'autres à vocation nationale. Pour résumer, les pôles de compétitivité représentent une nouvelle forme d'organisation territorialisée de l'innovation. Au sein de ces pôles,

spécifiquement les pôles à vocation mondiaux, se développent des réseaux ouverts à l'extérieur permettant aux agents d'être à la pointe de la technologie des partenaires différents. Ces réseaux aident donc à la transmission des connaissances entre les différentes firmes et entre les individus à travers les liens sociaux qui unissent les agents.

Chacune de ces approches et politiques territoriales d'innovation présentent le territoire comme "*acteur spécifique*" du processus de production de l'innovation. Contrairement à l'approche de la géographie de l'innovation, elles soulignent l'importance des interactions qui facilitent les transferts de connaissances, même si l'idée généralement défendue soutient que la proximité géographique facilite le transfert des connaissances, l'apprentissage et l'innovation.

1.4.1.2. Les effets de proximité dans les systèmes territoriaux d'innovation

Dans la continuité des travaux sur les districts industriels, ceux sur les systèmes productifs locaux (Becattini, 1990) ou sur les milieux innovateurs soulignent l'importance de la proximité géographique dans le processus de diffusion des connaissances. Même si l'importance de la proximité géographique semble être plus reconnue quand les connaissances tacites jouent un rôle essentiel dans la génération d'activités innovantes (Audretsch et Feldman, 1996), les autres formes de proximités non géographiques sont aussi déterminantes dans le processus de circulation des connaissances au sein des STP. La proximité géographique constituait une condition préalable, favorable à l'innovation. Les entreprises se localisent à proximité les unes des autres pour bénéficier d'externalités technologiques qui affectent positivement la fonction de production d'innovation. Elle a pourtant été relativisée par l'observation d'autres formes utiles de proximité, en prenant acte du fait qu'il n'est pas forcément nécessaire d'être localisé au sein d'un même espace pour innover.

Trois courants sont identifiés dans les travaux sur les différentes formes de proximités : interactionnistes, institutionnalistes et évolutionnistes (Balland, 2010). Le premier se focalise sur la capacité pour une organisation de faire interagir ses membres (Kirat et Lung, 1999). Le deuxième se positionne davantage sur l'étude des interactions elles mêmes que sur leur contexte, en distinguant proximité géographique et proximité organisée (Rallet et Torre, 2001). Le dernier fait référence à l'ensemble des routines implicites ou explicites facilitant la coordination entre les acteurs. Comme nous l'avons évoqué dans la section précédente, Boschma (2005) a proposé de différencier la proximité non géographique en quatre types de proximité : cognitive, organisationnelle, sociale et institutionnelle dans son approche évolutionniste qui présente l'avantage

de montrer pour chaque forme de proximité, comment un excès de proximité agit négativement sur l'innovation.

La proximité cognitive renvoie à la distance entre les bases de connaissance des acteurs. Elle est optimale quand les idées ou les connaissances sont complémentaires et similaires mais pas identiques. La proximité cognitive fait donc référence aux connaissances et aux expertises partagées entre acteurs. Pour certains auteurs, la proximité cognitive joue un rôle beaucoup plus décisif et complexe que la proximité géographique (Nooteboom, 2000). Dans une certaine mesure, la proximité cognitive augmente la capacité à capter la connaissance de l'autre (Brossard et Vicente 2010), et donc le potentiel bénéfique de la collaboration.

La proximité organisationnelle est liée aux logiques de coordination technico-productive. Elle caractérise par exemple la relation entre deux firmes d'un même groupe ou bien la relation entre un ensemble de fournisseurs et leur donneur d'ordres.

La proximité institutionnelle souligne quant à elle, l'intériorisation de normes, routines, règles, contrats, valeurs et langages communs. Elle se manifeste lorsque les acteurs produisent et partagent des représentations et des valeurs communes.

Enfin, *la proximité sociale* est l'encastrement des relations individuelles entre acteurs basées sur la confiance, l'amitié et les relations familiales. territoire est en effet avant toute chose un espace d'interactions et de lien social (Colletis, 2010).

Un territoire innovant ne bénéficie pas seulement de proximités géographiques mais aussi de proximités multiples pouvant contribuer à une meilleure circulation des connaissances.

1.4.2. Les canaux de transmission des connaissances

Les études théoriques et empiriques cherchant à expliquer l'inégale répartition des activités liées à l'innovation dans l'espace étudient la co-localisation des acteurs, sans considérer ni leurs liens, ni la structure relationnelle à laquelle ils appartiennent (Breschi et Lissoni, 2001). Ceci revient à masquer les vecteurs de transmission des connaissances notamment les réseaux de connaissances. En effet, même si l'effet de l'espace est démontré, nous ne pouvons pas conclure à l'effet crucial des *spillovers* sans étudier leurs canaux de transmissions.

Comme le mentionnent Krugman (1991a,b) et ensuite Breschi et Lissoni (2001), au delà de l'affirmation de l'existence de *spillovers* de connaissances, la question cruciale reste celle d'étudier les canaux de transmission des connaissances. En effet, l'étude des

canaux de transferts des connaissances constitue une problématique plus large que celle de la géographie de l'innovation.

Désormais, l'objectif central de ces approches récentes est d'expliquer l'organisation des échanges de connaissance au sein des territoires et de montrer comment différents réseaux contribuent à l'efficacité de la circulation des idées sur certains territoires. Parmi les différents éléments explicatifs relayés par la littérature sur les canaux de transmission des connaissances, sont mis en avant trois principaux canaux de diffusion : la mobilité des travailleurs, les spin-off et les réseaux sociaux. Les TIC, quant à elles, sont considérées comme un moyen facilitant le transfert de connaissances dans ces canaux de diffusion. Leur diffusion n'a donc pas provoqué une redistribution géographique des espaces innovants et des canaux de circulation des idées.

1.4.2.1. La mobilité des travailleurs

Les travailleurs ou les individus occupent une place importante dans la transmission des connaissances pour les deux raisons suivantes : d'une part, en raison du rôle des relations interindividuelles formalisées (Zucker et al., 1994), d'autre part, du fait de la mobilité inter-organisationnelle d'individus particuliers comme les ingénieurs ou les chercheurs (Almeida et Kogut, 1999, dans les semi-conducteurs et Stople, 2001, dans la technologie des cristaux liquides).

Plusieurs auteurs ont montré que les *spillovers* ne se diffusent pas seuls dans l'atmosphère mais qu'ils résultent de l'action d'acteurs particuliers comme les inventeurs (Zucker et al., 1998 et Lamoreaux et Sokoloff, 1997). La mobilité des inventeurs constitue alors une source d'externalités de connaissance particulière sur laquelle les entreprises ont dû mal à exercer un contrôle (Moen, 2000). Ces inventeurs représentent donc des individus clés "*intensifs*" dans la production de connaissances. Ils peuvent être des ingénieurs qualifiés ou des chercheurs qui transmettent leurs connaissances tacites dans les organisations qui les emploient successivement. Ce facteur joue un rôle important comme canal de diffusion des connaissances dans la plupart des travaux théoriques et empiriques sur les canaux de transmission des connaissances (Almeida et Kogut, 1999 et Breschi et Lissoni, 2009).

1.4.2.2. Le spin-off

Les essaimages (spin-off) constituent un deuxième canal de diffusion des connaissances. Klepper et Thompson (2010) proposent de différencier les sources de spin-offs en trois groupes. Un employé par exemple, pourrait faire une découverte d'une certaine valeur économique. La découverte est en principe valorisable par son entreprise. Cependant, les asymétries d'information d'une firme à une autre pourraient persuader

l'employé de mettre en oeuvre la découverte à travers sa propre start-up plutôt que de la révéler à son employeur. Dans la même lignée, Anton and Yao (1995), Wiggins (1995), Bankman et Gilson (1999), Gromb et Scharfstein (2002), Amador et Landier (2003), et Hellman (2007) précisent que d'une part, les entreprises ne peuvent pas s'engager dans un contrat qui récompense adéquatement l'employé pour sa découverte et d'autre part, les contrats de l'entreprise qui sont acceptés ex-ante, ne sont pas toujours suffisants pour empêcher un départ de l'employé. Cela constitue donc une première raison de création des spin-offs.

Dans le deuxième cas, la découverte est une connaissance commune dans l'entreprise mais elle est moins précieuse pour l'employeur que pour la start-up. En effet, l'implémentation de ces spin-offs représente un rival à l'entreprise puisque cette dernière possède moins de compétences pour évaluer les nouvelles compétences détenues par les employés, surtout quand l'idée touche les activités principales de l'entreprise¹⁵.

Dans le troisième cas, les employés apprennent de leurs employeurs comment réussir dans une situation de concurrence. Ils exploitent donc ce qu'ils ont appris comme connaissances pour créer leurs propres firmes dans la même industrie (Agarwal et al., 2004 ; Franco et Filson, 2006).

Le spin-off désigne donc une des techniques particulières de la scission, et même la société autonome qui en résulte. La société qui se sépare d'une partie de son activité est appelée société mère, et la filiale qui devient autonome suite au spin-off est définie par le terme société détachée ou société issue du spin-off. En fait, l'importance des spin-offs réside non seulement dans le fait qu'ils commercialisent les innovations mais aussi qu'ils favorisent la circulation des personnes qui ont un niveau très élevé de connaissances. Ainsi, certains employés quittent la société mère pour travailler dans les filiales. Ceci permet de transférer les connaissances des sociétés mères aux spin-offs¹⁶.

Les spin-offs ont joué un rôle important dans le développement de la Silicon-Valley et aussi ont permis le transfert à de longues distances, ce qui a contribué au développement de certains pays émergents. Pour cette raison, les essaimages ne doivent pas être négligés même si le rôle des travailleurs hautement qualifiés dans le domaine de l'innovation semble occuper une position centrale dans la création et la diffusion des connaissances (Glassmeier, 1988). Ils servent de catalyseurs dans le processus de transfert de connaissances et d'innovation technologique. Un troisième facteur de diffusion des connaissances a pris une place prépondérante dans la littérature et

15. Voir Pakes et Nitzan (1983), Tushman et Anderson (1986), Henderson et Clark (1990), Christensen (1993), Klepper et Sleeper (2005), et Cassiman et Ueda (2006).

16. Même si les spin-offs permettent le transfert de connaissances à longue distance, il existe tout de même des spin-offs localisés au près des filiales mères (Buenstorf, 2009).

plusieurs approches théoriques et empiriques se sont développées pour montrer son importance comme moyen de transfert de connaissances : les réseaux.

1.4.2.3. Les réseaux

Si l'approche de la nouvelle économie géographique a pris en compte l'importance des facteurs économiques dans la dynamique des STP (tels que la main d'oeuvre qualifiée, les coûts de transports et les externalités de connaissance localisées), une nouvelle approche s'est construite autour de la notion de réseaux (Powell, 1990 ; Conti et al., 1995) basée sur des facteurs territoriaux, sociaux, institutionnels et culturels (Rocha, 2004).

Le réseau est un ensemble d'entités sociales reliées entre elles par des liens créés lors d'interactions. Deux grands types de réseaux peuvent être distingués : les réseaux complets qui visent à appréhender l'ensemble des relations au sein d'une population donnée d'individus et les réseaux personnels où l'approche est centrée sur l'individu et sur les relations qu'il entretient (rapport DGCIS, 2011). Les agents au sein d'un réseau peuvent être de nature différente (chercheurs, firmes, clients, fournisseurs, concurrents mais également amis, etc.). Les objectifs de l'organisation en réseau peuvent être alors très variés selon qu'il s'agisse d'un réseau client-fournisseur, d'un réseau de chercheurs, etc.

Pour le cas d'un réseau d'innovation¹⁷, les sources d'innovation ne sont plus exclusivement au sein de l'entreprise mais fréquemment rencontrées aux frontières des firmes, des universités, des fournisseurs et des clients (Powell, 1990). Le lieu de l'innovation ne se situe plus dans l'entreprise mais au sein des réseaux de relations inter-organisationnelles (Powell, 1996)¹⁸.

Un réseau engendre une circulation des connaissances efficace si les membres sont suffisamment différents pour ne pas posséder les mêmes connaissances mais également suffisamment proches pour comprendre les connaissances qui circulent. De plus, un réseau est ouvert sur l'extérieur afin d'accéder à des partenaires différents. Il en résulte que les firmes peuvent appartenir à différents types de réseaux qui n'ont pas les mêmes caractéristiques. La coordination entre acteurs peut être facilitée via la mobilité des technologies des informations et de communication (TIC).

17. Nous précisons que l'étude s'intéresse seulement à une forme de réseau spécifique : les réseaux d'innovations. Les réseaux qui tiennent compte de l'ensemble des liens entre les noeuds, quelque soit la nature du noeud va au delà de l'objectif de ce travail.

18. Nous présentons dans la sous section 1.4.3 une analyse par les réseaux à travers les pôles de compétitivité.

1.4.2.4. Place des TIC dans le transfert des connaissances

Il est vrai que la proximité géographique joue un rôle important dans le transfert des connaissances dans la mesure où elle facilite les relations interindividuelles. Seulement, une première vision a remis en cause cette idée par la prise en compte d'autres types de proximités non spatiales que nous avons déjà précisé précédemment. Une autre vision affirme qu'avec l'avènement des Technologies d'Information et de Communication (TIC), les connaissances perdent leurs aspects géographiques. En effet, les TIC ont déjà exercé et continuent d'exercer un impact notable sur la transmission des connaissances.

Madeuf et al. (2005) ont identifié trois apports principaux des TIC sur les activités de R&D. Premièrement, la réduction de la part du savoir tacite à travers la codification des connaissances tacites et l'élargissement des capacités de compréhension en les rendant plus indépendantes du contexte. Certains outils de plus en plus utilisés, tels que les outils de travail collaboratifs ou la visioconférence pourraient permettre de transférer la connaissance tacite. Deuxièmement, l'informatisation démultiplie les capacités de traitement d'informations, ce qui élargit considérablement le rapprochement et le croisement de données initialement éloignées, favorisant ainsi le processus créatif. Finalement, les TIC peuvent développer la créativité, la conduite des projets, la facilitation des opérations de R&D et celle de la gestion de la propriété intellectuelle. Ainsi, nous pouvons considérer que les TIC permettent l'accélération des procédures et l'accroissement des quantités d'informations traitées. En effet, elles représentent un facteur d'accélération de l'innovation à travers la transmission des connaissances. Elles sont aussi un support à une production plus collective et plus interactive de la connaissance. Les travaux qui ont étudié la relation TIC-Connaissance sont plutôt d'ordre sociologique. Ils s'appuient sur des enquêtes pour expliquer l'effet des TIC sur le changement du comportement des individus dans les entreprises. Il en ressort généralement que les modifications ne sont pas importantes. En effet la densité des interactions s'est fortement accrue mais la distribution spatiale n'est pas modifiée. Dans la littérature économique, la question de l'impact des TIC sur la dynamique spatiale a été traitée de deux manières différentes. Un courant prônait l'abolition des distances grâce à la diffusion de technologies qui permettraient une décentralisation totale des activités (Thery, 1994 ; Cairncross, 1997). Les chercheurs en France peuvent communiquer avec des chercheurs aux Etats Unis ce qui est en faveur de la collaboration scientifique. Un autre courant, plus récent, affirme que les TIC renforcent les effets d'agglomération (Suire et Vicente, 2002).

Ainsi, selon Antonelli (1999), plus l'activité industrielle est dense dans la zone de localisation de la firme, plus son réseau de relations est dense et donc plus le niveau d'informations reçues sera important, ce qui augmente l'utilité d'adoption des TIC.

Avec le téléphone portable par exemple, les ingénieurs téléphonent plus souvent mais à des personnes proches géographiquement. L'internet est plus souvent utilisé en Intranet pour renforcer les liens locaux et est en général un peu verrouillé pour préserver la confidentialité interne. La coordination via les TIC s'accroît et permet de relier des unités situées dans des pôles urbains et qui bénéficient des externalités de proximité liées à leur localisation (Rallet et Torre, 2007). Il en résulte une économie d'archipels (Veltz, 2005) ou de flux (Castells, 2002), c'est-à-dire un ensemble de pôles urbains reliés par les réseaux de transport et de télécom. En même temps, la densité des relations tend à renforcer les phénomènes d'interdépendance et donc les phénomènes d'externalités de réseau (Katz et Shapiro, 1985). Ceci confirme la relation positive qui existe entre la densité d'une région et le degré d'adoption des technologies.

Il ressort de ces études que les TIC génèrent un effet densité (nombreux contacts) et un effet variété (facilité de trouver des informations) (Guillain et Huriot, 2000). Les TIC permettent donc l'éclatement spatial d'une organisation et de tirer profit de différentes externalités de connaissance. En effet, la diffusion des connaissances par l'intermédiaire des TIC semble illimitée géographiquement et ne dépend pas des nombres de frontières entre départements. En outre, la plupart des études empiriques qui ont été réalisées sur l'usage des TIC (Galliano et Roux, 2006 ; Soulié, 2006 ; Bocquet et Brossard, 2007 ; 2008) s'appuient sur des données désagrégées. Ces travaux étudient la relation entre l'usage des TIC par les entreprises et leurs localisations. Mesurer leurs effets dans le transfert des connaissances au niveau agrégé, notamment au niveau départemental, est une tâche difficile.

En revanche, il est difficile dans notre cas d'utiliser une mesure directe des TIC pour déterminer les frontières géographiques des connaissances.

1.4.3. De l'intérêt des pôles de compétitivité dans la transmission des connaissances : une analyse par les réseaux

La politique française des pôles de compétitivité s'inscrit dans la stratégie européenne définie lors du Conseil Européen de Lisbonne en Mars 2000 visant à faire de l'Union Européenne l'économie de la connaissance la plus compétitive et la plus dynamique du monde. Cette politique entamée par la France reposerait davantage sur une logique de mise en réseau des entreprises à l'échelle régionale, favorisant ainsi leur aptitude à innover.

Dans ces clusters interactifs en France (Cooke, 2001 ; Brette et Chappoz, 2007), le succès réside sur un ancrage territorial fort et sur l'importance de réseaux et de

liens plus étroits entre université et industrie. Les facteurs critiques de l'innovation sont inscrits dans une réalité sociale spatialement forte car leur conception même implique une approche territoriale. Néanmoins, ces pôles ne s'appuient pas seulement sur une proximité géographique ni sur des proximités sectorielles ou sur des compétences technologiques traditionnelles entre acteurs. La configuration actuelle des pôles montre en effet souvent une grande diversité d'appartenance sectorielle des différents membres d'un même pôle (rapport DGCIS, 2011). Dans les réseaux d'innovation, les flux de connaissance sont considérés à la fois comme une source d'incitation à coopérer et aussi comme la résultante des activités de coopération. Le phénomène de collaboration est au centre des analyses par ces réseaux d'innovation notamment les coopérations en R&D.

La variété des acteurs et le rôle spécifique du potentiel humain sont souvent déterminant dans la structuration des réseaux. En effet, la collaboration entre différents acteurs du réseau permet le développement des investissements en R&D grâce au partage des coûts, l'accès à des connaissances complémentaires et à l'existence d'externalités de connaissance.

1.4.3.1. Place de la coopération au sein des pôles de compétitivité

Zucker et al. (1994) sont les premiers à mettre en évidence économétriquement le rôle des interactions entre acteurs dans la diffusion des connaissances. Ils étudient les relations et la diffusion des externalités entre les scientifiques ne travaillant pas dans les entreprises mais qui ont co-publié avec les chercheurs de l'entreprise, les chercheurs purement académiques et les chercheurs affiliés dans les entreprises biologiques en Californie. Leurs résultats montrent que l'interaction entre scientifiques liés et les chercheurs affiliés à une entreprise ou avec des chercheurs purement académiques est bénéfique pour la diffusion des externalités de connaissance entre agents. Ces auteurs montrent clairement que la diffusion de *spillovers* n'aura lieu qu'en présence d'interactions et de liens entre agents. Ceci a été confirmé par d'autres travaux comme celui de Griliches (1998) qui affirme que : "*models of externalities must perforce be models of interaction between different actors in the economy*", (cité par Veugelers et DeBacker, 1999).

Une forme particulière d'interaction est présentée par les liens de coopérations. Cette forme a permis d'apporter quelques fondements théoriques à l'association spillover-coopération. Gallié (2003) précise que la décision de coopérer découle d'un arbitrage entre la maximisation des *spillovers* rentrants et la minimisation des *spillovers* sortants. Ceci est expliqué d'une part par le fait que les relations sont au moins partiellement marchandes et d'autre part par le fait que le courant dominant considère

que la coopération est un moyen d'internaliser les *spillovers*. Autrement dit, la circulation des *spillovers* dans les coopérations est liée à l'incomplétude des contrats. En effet, en raison de la non-rivalité, de la non-excluabilité et de la non cumulativité de la connaissance (même si l'on considère que la connaissance est parfaitement appropriable), il serait difficile d'en faire payer le coût exact dans de telles relations car l'usage potentiel de ce bien est incertain. Ainsi, il existe une incertitude forte sur l'usage privé que peuvent faire les cocontractants des connaissances qu'un des leurs a divulguées (Cohendet et Meyer-Krahmer, 2001). D'où le fait que la coopération peut être aussi une source de diffusion des connaissances liée aux interactions entre individus. Il découle de ceci que les externalités ont une place centrale dans les décisions de coopérer. En effet, les entreprises coopèrent non pas seulement pour partager les coûts et les risques et pour accroître leurs économies d'échelle mais aussi dans l'espoir de capter des externalités de connaissance de leurs partenaires (Mowery et al. 1997). Plus précisément, la diffusion des externalités de connaissance s'effectue de manière bi-latérale. En fonction de leurs capacités d'absorption, de leurs tailles et leurs avantages concurrentiels, ces firmes bénéficient de *spillovers* dont la quantité diffère selon les caractéristiques mentionnées. En revanche, ces firmes peuvent perdre une partie de leur avantage compétitif au cours de la coopération en raison des *spillovers* qu'elles émettent. C'est pourquoi, la décision de coopérer ou non résulte de l'arbitrage entre la maximisation des externalités rentrantes et celui de minimisation des externalités sortantes (Gallié, 2002).

1.4.3.2. Impact de la coopération sur la diffusion des connaissances et sur l'innovation

Dans la littérature sur les districts industriels, les réseaux d'innovation, les milieux innovateurs et plus récemment sur les pôles de compétitivité, les flux de connaissance peuvent se produire d'une manière volontaire et involontaire entre organisations. En effet, les accords de coopération permettent d'internaliser les *spillovers* de connaissances ce qui donne des incitations supplémentaires aux firmes pour investir en R&D (Czarnitzki et al., 2007; Karray, 2001). Cet investissement en R&D, à son tour, a un effet positif sur la production d'innovations. Plusieurs études se sont développées afin de mesurer l'effet des flux de connaissance et l'effet de la R&D sur l'innovation (Iammarino et al. 2009, Czarnitzki et al. 2007, Becker et Dietz 2004, Belderbos). En utilisant l'enquête communautaire sur l'innovation (CIS), ces auteurs montrent que les entreprises qui coopèrent sont plus fortement engagées dans des activités innovantes que les entreprises qui ne sont pas impliquées dans des accords de coopération en R&D, et ce quelque soit la définition retenue pour mesurer la performance innovante (introduction effective d'innovations, ventes associées aux nouveaux produits, dépôts de brevets) et la zone géographique considérée.

1.4.3.3. Capacité d'absorption et succès de collaboration

La coopération constitue un des canaux de diffusions des *spillovers* dans la mesure où elle favorise les interactions entre partenaires. Son succès dépend de plusieurs caractéristiques dont la principale est la capacité d'absorption (Gallié, 2003). Cette capacité d'absorption peut être divisée en trois catégories. La première catégorie est la capacité de production ou la capacité de faire fonctionner des installations. Cette capacité inclut le management de la production, l'ingénierie de production, la réparation et la maintenance du capital physique. La deuxième représente la capacité d'investissement : c'est-à-dire la capacité d'entreprendre des activités d'identification, préparation, conception, aménagement et instruction de nouveaux projets. Enfin, la capacité d'invention qui englobe la compétence de l'entreprise à adapter, améliorer et développer la technologie. Cohen et Levinthal (1989) ont mis en évidence la nécessité de construire une capacité d'absorption afin de bénéficier des connaissances révélées par d'autres. Selon ces auteurs, en présence d'externalités de connaissance fortes, les entreprises peuvent avoir intérêt à augmenter leurs investissements en R&D afin de construire une capacité d'absorption et ainsi de bénéficier plus efficacement des connaissances externes : ceci contribue à créer un lien positif entre externalités de connaissance et incitations à la recherche.

1.5. Conclusion

Jusqu'aux années 1980, le thème de l'innovation est souvent délaissé et les premiers travaux qui sont apparus se sont focalisés sur l'étude de ses déterminants traditionnels tels que le volume de la R&D. Les travaux évolutionnistes montrent que l'innovation est liée aux connaissances qui constituent l'une de ses variables déterminantes. La connaissance est un bien public, mais en partie : elle ne suscite pas de rivalité puisque plusieurs utilisateurs peuvent l'utiliser en même temps mais elle est non exclusive en partie dans la mesure où son propriétaire ne peut empêcher autrui de l'utiliser que dans une certaine mesure. Il en résulte des bénéfices pour les firmes externes qui enrichissent leurs connaissances initiales, ces bénéfices sont alors des externalités de connaissance. Même si la littérature sur la notion d'externalités de connaissance est éparse, leur rôle sur les activités innovantes est important.

Dans ce chapitre, nous ne sommes pas intéressés à l'étude de la différence entre flux de connaissance et externalités de connaissance. Nous avons essayé plutôt de concentrer notre problématique sur l'étendue géographique des différents types de

flux et connaissances leurs canaux de transmissions¹⁹. Pour cela, nous avons mobilisé d'une part, la littérature en économie géographique, économie de la connaissance et géographie de l'innovation pour expliquer le caractère localisé (Localised Knowledge) et délocalisé (Distant Knowledge) de la connaissance. D'autre part, nous avons montré à travers la présentation de différents canaux de diffusion des connaissances que si l'espace compte dans le transfert des connaissances, c'est parce qu'il existe des réseaux d'innovation qui facilitent la collaboration et qui permettent de créer et de diffuser la connaissance, et que ces réseaux sont souvent localisés, même si certains d'entre eux peuvent s'établir sur de plus longues distances.

Nous proposons dans le chapitre suivant de présenter les différentes méthodes de mesure des *spillovers* de connaissances notamment la fonction de production de connaissance ou d'innovation (FPI). En utilisant des données départementales, nous tentons ensuite de mesurer l'étendue géographique des connaissances et leurs effets sur l'innovation au niveau de la France métropolitaine.

19. Nous avons donc tendance par la suite à pratiquer le même abus de langage que la plupart des autres études qui confondent, à l'intérieur du vocable commode "*spillovers*", les flux de connaissance et les externalités de connaissance

Chapitre 2

Les outils de mesure de la diffusion spatiale des connaissances : Quelle est la bonne approche empirique ?

2.1. Introduction

Dans les modèles de croissance endogène, la dynamique économique est principalement causée par la création des connaissances et le progrès technologique (Romer, 1986 ; 1990). Ainsi la croissance économique dépend, en plus des activités d'innovation locales, des connaissances externes fournies par les externalités de recherche et développement (Grossman et Helpman, 1991 ; Coe et Helpman, 1995 ; Paci et Usai, 2000 ; Bottazzi et Peri, 2003). Ces externalités ont pris une place prépondérante dans la littérature récente sur les facteurs d'innovation.

De larges travaux empiriques se sont penchés sur l'étude des externalités via les mécanismes de diffusion des connaissances. Ces travaux empiriques peuvent être classés en deux groupes. Un premier groupe met en évidence deux types d'externalités : les externalités de spécialisation qui fonctionnent principalement au sein d'une industrie spécifique, et les externalités de diversification qui favorisent la création de nouvelles idées entre les secteurs (Paci et Usai, 2000). Un deuxième groupe met en avant la différence entre l'aspect public et l'aspect privé des connaissances. Les études n'aboutissent pas toujours aux mêmes conclusions : Coe et Helpman (1995) à titre d'exemple affirment que le progrès technique est d'intérêt public et par conséquent, la diffusion des connaissances peut librement se faire à travers les frontières. Par opposition à cette idée, d'autres études ont montré que les connaissances sont de nature privée et locale (Jaffe et al., 1993). Selon ces derniers auteurs, l'acquisition de

la connaissance est un processus long et coûteux. Par conséquent, il est difficile de les transmettre entre les secteurs d'où l'intérêt de la proximité spatiale dans ce processus de partage et de diffusion des connaissances.

A partir de ces études, nous constatons qu'il existe un consensus assez large quant à l'existence des externalités de connaissance même si elles sont inobservables. En effet, il est difficile d'identifier leur ampleur et leur étendue géographique. La multiplication des travaux empiriques étudiant les externalités de connaissance témoignent de cet état de fait. Cependant, comme nous l'avons indiqué dans le chapitre précédent, ces travaux emploient excessivement le terme externalités de connaissance pour désigner la diffusion ou les flux de connaissance. Par ailleurs, ces travaux présentent plusieurs apports en termes de mesure des externalités de connaissance et de leurs effets sur l'innovation.

Dès lors, l'objectif de ce chapitre est double. Dans un premier temps, nous essayons de mettre en lumière les principaux modèles empiriques cherchant à étudier la relation entre innovation et externalités de connaissance. Nous résumons les travaux effectués dans deux approches : l'approche par les citations de brevets et l'approche par la Fonction de Production de Connaissance (Knowledge Production Function ou KPF). Dans un deuxième temps, nous tentons d'appliquer l'approche par cercles concentriques de Bottazzi et Peri (2003) sur un échantillon de données françaises.

Notre chapitre 2 est organisé comme suit. La section 2 décrit l'approche par les citations de brevets qui constitue un marqueur des externalités. La section 3 présente l'approche par la Fonction de Production de Connaissance (KPF). La section 4 s'intéresse au cadre d'analyse empirique et aux données. Les sections 5 et 6 présentent l'étude en coupe et le panel étudié. Enfin, la section 7 conclut le chapitre.

2.2. L'approche par les citations de brevets

Dans la littérature, les citations de brevets sont généralement utilisées pour deux buts. Le premier consiste à apprécier la qualité d'un brevet, laquelle augmente avec le nombre de citations reçues. Cela permet d'évaluer la valeur économique du portefeuille de brevets des entreprises. Le deuxième objectif de l'utilisation des citations de brevets consiste à mettre en évidence des traces directes de la diffusion des connaissances ou "*marqueurs de spillovers*". Ce deuxième objectif est crucial pour les économistes qui cherchent à prouver l'existence des externalités de connaissance. Les citations sont considérées comme des traces écrites d'un transfert des connaissances de l'inventeur ou de l'applicant cité vers l'inventeur ou l'applicant qui cite. Plus précisément, les citations de brevets sont la partie des informations contenues dans les brevets. Chaque

brevet comporte des références qui répertorient l'ensemble des brevets décrivant l'état de l'art associé à des technologies antérieures.

Cette approche présente l'avantage d'apporter des évidences empiriques sur les flux de connaissance entre inventeurs ainsi que sur leur trajectoire de diffusion : les brevets originels ou encore les brevets cités sont des brevets qui ont servi de référence pour un nouveau brevet déposé. Ces citations permettent d'approximer les flux de connaissance et d'appréhender leur géographie. Les citations de brevets ont été proposées par Jaffe, Trajtenberg et Henderson (JTH) (1993) en réponse à Krugman (1991a) selon lequel les économistes devraient abandonner toute idée de mesurer les externalités de connaissance. A l'inverse de Krugman (1991a), JTH considèrent que les externalités technologiques laissent des traces : les citations de brevets. Cette marque laissée par l'innovation peut permettre d'identifier, au moins partiellement, la trajectoire de diffusion des connaissances et leur caractère géographique. L'analyse de ces auteurs consiste à comparer la localisation géographique des citations avec celle du brevet originel. Cette méthode ensuite reprise par Jaffe et Trajtenberg (1999), Almeida et Kogut (1999), en confirmant les résultats de Jaffe et al., (1993) a montré que la probabilité pour une citation d'être localisée dans la même zone géographique que le brevet originel est très largement supérieure à la probabilité pour n'importe quel brevet d'être localisé dans cette même zone.

2.2.1. Principe sous-jacent

Le système de brevet prévoit deux sortes d'indications sur les connaissances préalables qui accompagnent une innovation : les références à la littérature scientifique et les références à d'autres brevets ou citations de brevets. Elles sont les plus utilisées dans la mesure où elles fournissent des indications précises sur l'innovateur et sur sa localisation. Elles consistent à compter le nombre de citations reçues par brevet publié (Trajtenberg, 1990 ; Hall et Ziedonis, 2001). Autrement dit, le nombre de fois qu'un document de brevet est cité dans d'autres documents de brevets. Ces citations ont été utilisées comme une technique afin d'estimer la valeur économique d'un brevet (Harhoff et al. 1999). Ainsi, les résultats des deux études menées par Harhoff et al. (1999, p511), en Allemagne et aux États-Unis, indiquent une relation positive entre la fréquence des citations et la valeur économique déclarée. Ces résultats confirment ainsi l'influence des connaissances intégrées à des inventions antérieures sur l'émergence de nouvelles. Plus encore, selon Hall et al. (2001), les citations de brevets sont économiquement utiles puisqu'elles ne demandent guère d'investissement en R&D. En plus de son rôle de mesure de la valeur économique, les citations de brevets jouent un rôle primordial dans le repérage des externalités de connaissance et notamment de leur dimension

géographique. Elles représentent la trace de la diffusion des connaissances et en traduisent le flux. Si un brevet B cite un brevet A, cela implique que le brevet A représente une partie de la connaissance déjà existante, sur laquelle s'appuie le brevet B. Par conséquent, le nombre de citations reçues pour chaque brevet nous renseigne sur les quantités d'informations transmises et constitue de ce fait un indicateur des externalités de connaissance. Un premier groupe de travaux met en évidence une coïncidence de la localisation des flux de connaissance et de leurs sources (§2.2.2), alors qu'un second groupe teste l'impact de l'éloignement géographique sur l'intensité des flux de connaissance (§2.2.3).

2.2.2. Co-localisation entre brevets originels et citations

Jaffe, Trajtenberg et Henderson (JTH) (1993) sont les pionniers dans la mesure directe des externalités locales de connaissance. Leur expérience consiste à comparer la localisation géographique des citations avec celle du brevet originaire. Ces auteurs considèrent que les flux de connaissance laissent des traces dans la mesure où tout demandeur d'un nouveau brevet, au moment de sa demande de dépôt de brevet auprès de l'office compétent, doit mentionner dans le document qu'il soumet les connaissances antérieures (brevets, publications scientifiques, etc.) sur lesquelles se base son propre travail. Les relations entre brevets originels et brevets subséquents y faisant référence sont ainsi utilisées pour approximer les flux de connaissance. Afin de mener cette analyse en terme de citations de brevets, JTH (1993) prennent en compte le fait que les activités technologiques peuvent être au préalable concentrées géographiquement : "*The most difficult problem confronted by the effort to test for spillover localizations the difficulty of separating spillovers from correlations that may be due to a pre-existing pattern of geographic concentration of technologically related activities*". Ce problème d'auto-corrélation spatiale s'explique à la base par une concentration spatiale préexistante des activités technologiques dans un endroit donné. Ce qui fait que la coïncidence géographique entre deux brevets peut s'expliquer par la concentration des inventeurs au même endroit et non pas par la diffusion des connaissances. En effet, sans que ceci traduise l'existence de *spillovers*, nous avons plus de chance d'avoir des brevets provenant de zones géographiques fortement innovantes que d'ailleurs. En effet, si 20% des brevets proviennent de la Silicon Valley, il y a 20% de chance pour qu'un brevet situé dans la Silicon Valley cite un brevet de cette même région, indépendamment de l'existence des *spillovers*. Afin de contourner ce problème, Jaffe et al. (1993) vont donc construire un échantillon de brevets de contrôle. L'objectif c'est de comparer la probabilité qu'un brevet de contrôle appartenant à une même classe technologique et ayant la même date d'application, coïncide géographiquement

avec le brevet originel, sous la condition qu'il n'existe aucun lien de citation entre brevet de contrôle et brevet originel. Plus précisément, ces auteurs testent l'hypothèse nulle, où il n'y a pas de localisation des externalités de connaissance :



Figure 2.1 – Comparaison entre brevets de contrôle (B.Cr), brevets originels (B.O) et citations de brevets (B.Cit)

$H_0 : P_{BCit} = P_{Bcr}$ et l'hypothèse alternative

$H_1 : P_{BCit} > P_{Bcr}$,

où P_{BCit} représente la probabilité que le brevet qui fait la citation soit localisé au même endroit que le brevet originel et P_{Bcr} , la probabilité que le brevet de contrôle soit localisé au même endroit que le brevet originel.

Si la fréquence de la coïncidence géographique entre deux brevets liés par des citations est plus importante que celle des deux brevets non liés (dont nous avons contrôlé le temps et la technologie), alors nous pouvons dire que les citations sont activées au niveau local. Cette méthode a été largement utilisée dans les travaux mesurant les externalités de connaissance. Cependant, elle se cantonne à tester seulement la coïncidence géographique entre émetteur et récepteur de flux de connaissance sans mesurer l'éloignement géographique entre acteurs et son impact sur ces flux. En réponse à cette critique, un deuxième type de travaux utilisant les citations de brevets a été développé. En se basant sur des données américaines, Jaffe et al., (1993) montrent que la probabilité qu'une citation soit localisée dans la même zone géographique que le brevet originel, est largement supérieure à la probabilité que n'importe quel brevet soit aussi localisé dans cette même zone. Ils en déduisent alors que les externalités de connaissance sont géographiquement contraintes. Cette méthode a été ensuite reprise par Jaffe et Trajtenberg (1999), Almeida et Kogut (1999) confirmant ainsi les résultats de Jaffe et al. (1993). Almeida et Kogut (1999) élargissent par la suite leurs analyses à plusieurs pays et se focalisent sur la diffusion des connaissances entre les pays membres du G5 : la France, l'Allemagne, la Grande-Bretagne, les Etats-Unis et le Japon. Ces

auteurs s'attachent à expliquer la fréquence des citations entre brevets issus de paires de pays différents et également à comprendre son évolution au fil du temps. Les résultats économétriques trouvés confirment ceux de Jaffe et al. (1993) dans la mesure où ils démontrent une forte tendance à la citation au sein du même pays émetteur et que les citations sont principalement le fruit d'acteurs nationaux. Cependant, la concentration géographique des citations s'estompent avec le temps.

2.2.2.1. L'approche de Jaffe, Trajtenberg et Henderson (approche JTH) (1993)

L'étude de Jaffe, Trajtenberg et Henderson (1993) porte sur deux cohortes de brevets : l'une pour 1975 (950 brevets qui ont donné lieu à 4750 citations jusqu'en 1989) et l'autre pour 1980 (1450 brevets qui ont donné lieu à 5200 citations pour la même période). Chaque cohorte comporte un échantillon de brevets déposés par les universités américaines et un échantillon de brevets déposés par d'autres organismes. L'objectif est d'observer si les *spillovers* varient en fonction de l'institution qui se trouve à l'origine du brevet. Dans certains cas, la citation de brevet et le brevet d'origine sont déposés par le même organisme : c'est ce que nous appelons auto-citation. Ces auto-citations sont liées à l'existence d'un stock des connaissances préalable qui reste interne à la firme ou à l'organisme qui a déposé un brevet d'origine ; il ne s'agit pas donc de *spillovers*. C'est pour cette raison que JTH (1993) éliminent de leurs résultats les auto-citations. La localisation est étudiée à trois niveaux : le pays (Etats-Unis ou reste du monde), les Etats américains et les aires métropolitaines. Ces auteurs testent l'hypothèse selon laquelle les brevets qui citent les brevets d'origine américaine sont plutôt localisés aux Etats Unis. Le même travail a été mené au niveau des Etats et des aires métropolitaines en passant par l'indice d'Herfindall¹. Les résultats trouvés par JTH (1993) montrent que les citations de brevets déposés en 1980 et celles déposés en 1975 ont tendance à être localisés dans la même zone géographique que les brevets cités, mais avec une tendance moins prononcée pour les citations de brevets déposés en 1975.

Néanmoins, pour JTH (1993) la proximité technologique ne semble pas jouer de rôle particulier. C'est dans leur *working paper* de 1992 que les auteurs se sont intéressés à la dimension technologique des externalités. Sur la base des classifications officielles des brevets par secteur d'activité, les auteurs ont distingué environ 300 domaines de brevets primaires. Il semble donc que les brevets cités et les brevets qui les citent sont technologiquement proches. Mais les observations empiriques de l'article de 1993

1. Indique dans quelle mesure les brevets qui font référence à un autre brevet se distinguent les uns des autres en termes de localisation technologique.

contredisent ce résultat et indiquent qu'une part significative des *spillovers* provient de firmes qui ne sont pas dans le voisinage technologique immédiat de la firme considérée. Une autre dimension a été prise en compte dans les travaux de JTH (1993) : c'est la dimension temporelle des *spillovers*. JTH (1993) constatent en effet que pour qu'une connaissance incorporée dans le brevet soit suffisamment diffusée, il faut quelques années ou au moins quelques mois entre le moment où le brevet initial est déposé et le moment où ce brevet est cité. Par ailleurs, les auteurs observent effectivement que la correspondance géographique entre les brevets cités et les brevets qui les citent décroît au fur et à mesure que le temps passe et que les connaissances se diffusent. En outre, ils ont distingué les citations en provenance des universités de celles faites par d'autres organismes privés. Ils ont trouvé que ces derniers sont plus localisés que celles des universités. Thomson et Fox-Kean (TFK) (2005) testent le modèle de JTH (1993) en construisant deux groupes de contrôle. Dans le premier groupe, à chaque brevet citant, ils associent un brevet-contrôle ayant une date d'application similaire et la même classe technologique. Le second groupe sélectionne les brevets-contrôle ayant la même classification principale (au niveau des sous-classes) que les brevets-citant. Cette nouvelle base de données a été sélectionnée à travers différents critères : un niveau de classification technologique plus désagrégé et la classe de brevets original-citant-contrôle ayant au moins une classe en commun. Les auteurs trouvent que seule l'hypothèse de localisation des externalités au niveau national survit à la réévaluation du travail de JTH (1993). Ils ne trouvent pas par contre une évidence empirique quant à la localisation au niveau des états. JHT (2005), dans leur article "*Patent Citations and the Geography of Knowledge Spillovers : A Reassessment Comment*", ont répondu aux critiques de TFK (2005) en précisant que les critères de sélection adoptés par ces derniers font que le nombre d'observations a sensiblement diminué, biaisant en quelque sorte l'échantillon.

JHT (2005) ont donné comme exemple, l'étude faite par le professeur Robert Langer à l'Institut Technologique du Massachusetts, résumant ainsi l'essentiel de l'approche basée sur les citations de brevets². Le professeur Robert Langer est l'un des experts mondiaux dans l'ingénierie tissulaire. Il a déposé 120 brevets dans la région de Boston. Une grande partie des citations des brevets coïncident géographiquement. Les brevets du Professeur Langer ont reçu 1264 citations. Les brevets citant sont signés par 3553 inventeurs. 982 parmi ces 3553 inventeurs, soit 28%, résident à Boston. Les 72% restant sont dispersés sur plus de 500 différentes localisations. Bien que 28% des inventeurs citant résident à Boston, il nous paraît insatisfaisant d'admettre l'existence

2. Voir l'article de Henderson, Jaffe et Trajtenberg (2005) pour plus de détails coïncident géographiquement.

des externalités locales de manière évidente et directe. Comme nous l'avons précisé dans le chapitre premier, il existe d'autres facteurs d'agglomération que les externalités de connaissance. La spécialisation de la ville dans l'ingénierie tissulaire figure parmi ces facteurs qui peuvent expliquer l'agglomération des inventeurs à Boston. En effet, Boston est une ville mondialement connue pour ses compétences dans ce domaine ce qui peut expliquer la co-localisation des firmes spécialisées et donc la tendance des gens à habiter dans la même région. Après avoir exposé les principaux apports des travaux initiateurs de Jaffe, Trajtenberg et Henderson (1993) nous présentons dans ce qui suit les résultats majeurs des principales études qui ont suivi leur travail pionnier.

2.2.2.2. Approche d'Almeida et Kogut (1997a,b ; 1999)

Les travaux de Almeida et Kogut (1997a,b ; 1999) se focalise sur l'étude de la relation entre la taille des firmes et leur capacité d'innovation. Cette étude est différente de celle de Jaffe et al. (1993) : les auteurs ont comparé les domaines d'innovation des grandes firmes avec celles des petites firmes pour le secteur des semi-conducteurs et montrent que les petites firmes explorent des champs technologiques plus variés. Plus précisément, les petites firmes innoveraient davantage car elles appartiendraient à des réseaux locaux. A l'instar de Jaffe et al.(1993), les auteurs ont construit un échantillon de brevets de contrôle en utilisant les brevets déposés à la fois par les firmes américaines et par les firmes étrangères, localisées aux Etats-Unis. Leurs résultats confirment l'existence de facteurs externes autres que les *spillovers* dans la production d'innovation.

Almeida et Kogut (1997a,b ; 1999) constatent également que pour le cas de l'industrie des semi-conducteurs, les citations sont localisées, en cela, la dimension géographique est importante. L'apport de l'analyse d'Almeida et Kogut (1999) consiste à décomposer les résultats par région, et à permettre ainsi des comparaisons. En constituant trois échantillons pour 1975, 1980 et 1985, ils étudient les citations de ces brevets sur les huit années qui suivent leur dépôt. Ils trouvent une grande disparité dans les degrés de localisation. Le cas le plus illustratif est la Silicon Valley qui a en particulier un effet de localisation très fort. Selon Almeida et Kogut (1997b), le caractère localisé des externalités de connaissance résulte de la dimension tacite des connaissances. D'où le rôle de la mobilité de la main d'oeuvre dans le transfert des connaissances d'une firme à une autre.

Pour tester ces hypothèses, Almeida et Kogut (1997b) repèrent les trajectoires professionnelles des détenteurs de brevets. Ces trajectoires sont reconstituées à l'aide des brevets déposés entre 1974 et 1994 (les brevets indiquant le nom de l'ingénieur à l'origine de l'innovation et la firme dans laquelle il ou elle travaille). Deux résultats ont été tirés de cette étude : une grande disparité régionale dans le degré de mobilité inter

firmes de la main d'oeuvre et le deuxième résultat montre que la mobilité inter-régionale est très faible, ce qui explique la dimension locale des *spillovers*. Pour confirmer leurs résultats, les auteurs ont construit par la suite un nouvel échantillon de contrôle : l'objectif est d'observer si les firmes qui embauchent un nouvel ingénieur, détenteur d'un brevet, ont plus de chance de citer le brevet de cet ingénieur. Les résultats de ce test sont significativement positifs : les firmes qui embauchent un nouvel ingénieur tendent plus à citer le brevet de cet ingénieur que les autres firmes.

A travers cette étude, Almeida et Kogut (1997b) parviennent ainsi à apporter une explication à la question des modalités de transmission des *spillovers*. La mobilité du travail joue aussi un rôle important dans le transfert des connaissances. Toutefois, ces travaux sont insuffisants puisqu'ils se cantonnent seulement à tester la localisation géographique des citations avec celle du brevet originel cité et n'offrent pas la possibilité de mesurer l'impact de l'éloignement géographique entre acteurs. Pour cela, un autre type de travaux se développe ayant pour objectif de tester l'impact de la distance géographique entre entités (acteurs, territoires) sur l'intensité des flux de connaissance approchés par les citations de brevets.

2.2.3. Proximité géographique et flux de connaissance

En utilisant un échantillon de PME suédoises, Sjöholm (1996) utilise les citations de brevets comme indicateur de flux de connaissance de l'année 1986. Il étudie l'impact de la proximité géographique et du commerce international sur la probabilité du transfert des connaissances entre la Suède et les autres pays du monde. En utilisant une modélisation logistique, Sjöholm (1996) s'efforce d'identifier les déterminants de ces citations. L'auteur trouve que l'effet bénéfique de la proximité géographique (mesurée par la distance en milliers de kilomètres entre le centre économique Suédois et le centre économique des pays cités d'après les informations de l'US Naval Océanographique office) est remis en cause si d'autres variables explicatives sont incluses dans le modèle. Maurseth et Verspagen (2002) étudient l'intensité des flux de connaissance entre régions européennes en les approximant par les citations de brevets. Leurs résultats montrent que la distance géographique a un impact négatif ce qui explique que les flux de connaissance sont plus nombreux d'une part entre régions proches, et d'autre part au sein des pays qu'entre régions de pays différents.

En ce qui concerne la proximité technologique, Jaffe et al. (1993) et Verspagen et Schoenmakers (2000) montrent que la probabilité de citer un brevet est positivement influencée par la proximité technologique entre le nouveau brevet et le brevet cité (Massard et Riou, 2002). Plus récemment, Abdelmoula et Bresson (2007) mesurent l'impact de la proximité sectorielle sur le nombre de citations de brevets. L'objectif

principal de cette étude est d'analyser un réseau social d'entreprises formé par les citations de brevets entre elles. En effet, les auteurs étudient les caractéristiques de ce réseau social en observant en quoi la place de l'entreprise dans ce réseau peut influencer le nombre de citations que reçoivent ses brevets. A travers une base de données sur les citations de brevets européens, les auteurs ne considèrent que les citations extérieures reçues ; c'est-à-dire les citations reçues par des entreprises autres que celle qui a fait son dépôt tout en excluant les autocitations.

Dans un premier temps, les auteurs montrent à travers un réseau formé par 361 entreprises européennes (appartenant à 8 pays) que ce réseau est un petit monde dont la formation est basée sur l'attachement préférentiel. Dans une deuxième partie, ils estiment un modèle de comptage des citations de brevets et montrent, entre autres, que la centralité de l'entreprise déposante dans le réseau social influence le nombre de citations que reçoivent ces brevets. Néanmoins, les limites attribuées à l'approche par les citations de brevets sont multiples (Breschi et Lissoni, 2004).

2.2.4. Limites de l'approche par les citations de brevets

L'approche citation de brevets fondée par Jaffe et al.(1993), reprise par la suite par Jaffe et Trajtenberg (1999) ; Almeida et Kogut (1999) a été critiquée sur plusieurs points. En plus du fait que les critiques qui s'appliquent aux brevets, par conséquent, toutes les critiques qui s'appliquent aux brevets s'appliquent aussi aux citations de brevets³, nous présentons dans ce qui suit différentes critiques qui ont été adressées aux citations de brevets : niveau trop agrégé des classes de brevets (§ 2.4.1.1), les citations ne traduisent pas toujours des *spillovers* dans la mesure où les inventeurs ont la possibilité de citer des brevets après avoir terminé l'invention (§ 2.4.1.2), difficultés de détermination de la localisation de la citation (§ 2.4.1.3).

2.2.4.1. Niveau trop agrégé des classes de brevets

La méthode JTH a été critiquée par Thompson et Fox-Kean (TFK)(2005). Les résultats de ces derniers contredisent ceux trouvés par JTH (1993). Précisément, leurs critiques concernent l'échantillon des brevets de contrôles utilisés. Selon TFK (2005), les brevets de contrôles utilisés par JTH sont sélectionnés en se basant sur une nomenclature de classification des activités technologiques trop agrégée (3 niveaux) de l'USPTO. Ce niveau supprime un certain nombre d'hétérogénéité intra-classe puisque les brevets sont affectés à des classes de manière parfois arbitraire. En conséquence, ceci peut expliquer la différence entre les brevets de contrôle, les brevets originels et les

3. Puisqu'il y a un lien de citation entre un brevet originel et une citation de brevet

brevets citants(ou citations), ce qui garantit de plus cette similarité industrielle entre ces différents types de brevets.

2.2.4.2. Les citations ne traduisent pas toujours des *spillovers*

Thomson (2006) soulève d'autres biais liés à l'utilisation des citations de brevets dans la mesure des *spillovers*. En fait, les inventeurs ont la possibilité de citer des brevets après avoir terminé l'invention. Dans ce cas, cette partie des connaissances ne reflète pas un flux réel des connaissances. Un autre biais fait référence aux examinateurs qui rajoutent souvent des citations de brevets, non-mentionnés par les inventeurs (par oubli ou méconnaissance) alors qu'ils les jugent pertinents. Ceci peut être la cause de la surestimation de l'ampleur de la coïncidence géographique entre un brevet citant et un brevet originel.

De plus, au moment de son invention, l'inventeur n'avait pas nécessairement connaissance du brevet cité, et même s'il en avait connaissance, il ne s'en est pas forcément servi pour son invention. Ainsi, les citations de brevets ne traduisent pas forcément des relations effectives entre les nouveaux brevets et les brevets auxquels ils font référence.

Pour confirmer ces résultats, une étude a été faite par Jaffe et al. (2000) sur un échantillon de près de 2000 inventeurs. Leur objectif était de repérer si les citations correspondaient ou non à des *spillovers*. Pour cela, ils interrogent les inventeurs sur les liens éventuels qu'ils ont eu avec les déposants à l'origine des brevets qu'ils citent. Ils ont trouvé qu'à peu près la moitié des citations ne correspondrait pas à des flux de connaissance. De ce fait, une citation ne représente pas forcément un flux, puisqu'une partie des références citées dans les brevets déposés reste inconnue de l'inventeur. Pour cette raison, "*This makes patent citations a noisy measure of the extent and direction of the knowledge flows*" (Giuri et al., 2007). En outre, une externalité peut exister sans aucune citation de brevets. En effet, une innovation qui est mesurée par un autre indicateur que le brevet (articles scientifiques, publications, etc.) peut produire des externalités de connaissance.

2.2.4.3. Difficultés de détermination de la localisation

JTH (1992, 1993) ont sans doute déterminé trop arbitrairement la localisation des innovations. Les brevets peuvent fournir des indications précises sur le pays d'origine de l'inventeur. Néanmoins, une innovation peut avoir plusieurs inventeurs, qui n'habitent pas forcément au même endroit. Dans ce cas, les auteurs ont attribué arbitrairement en plusieurs lieux différents une même innovation, en la rattachant par exemple à l'Etat d'origine de l'un des inventeurs et à l'aire métropolitaine de l'autre inventeur. Ce cas

d'inventeurs multiples peut donc biaiser les résultats. Il faut toutefois remarquer que les cas d'inventeurs multiples ayant une localisation différente sont marginaux. Au niveau des pays, seuls 2% des brevets ont des inventeurs dans deux pays différents. Pour les Etats, 10% des brevets sont déposés par des inventeurs issus d'Etats distincts, et pour les aires métropolitaines, le problème se pose dans 14% des cas.

Toutefois, la question des effets bénéfiques de la proximité géographique a été remise en question. JTH (1992; 1993) ont trouvé que la proximité géographique est très sensible à toute modification des variables explicatives incluses dans la modélisation. Leur étude a été par la suite complétée par l'introduction d'un indice de proximité entre les structures industrielles des pays. L'objectif est d'analyser l'étendue des brevets suédois par des multinationales. Les auteurs ont trouvé un impact négatif mais significatif de la distance géographique sur la diffusion internationale des flux de connaissance. L'apport de ces études réside dans la précision que la seule proximité géographique n'explique pas à elle seule l'existence d'externalités de connaissance, mais que d'autres variables (telles que la taille de la firme) peuvent avoir un effet globalement plus intéressant dans la géographie de l'innovation.

Moreno et al (2005), Le Sage et al. (2007) et Paci et Usai (2009) adoptent la même méthodologie. Ils trouvent grâce à une analyse économétrique que la distance géographique constitue un obstacle pour la circulation des connaissances. Par ailleurs, les régions qui ont des frontières communes sont plus susceptibles de citer mutuellement leurs brevets respectifs. Ils ont montré aussi que les frontières nationales constituent un obstacle aux flux de connaissance dans la mesure où les citations de brevets sont plus élevées lorsque les deux régions appartiennent au même pays.

Bien que les travaux utilisant les citations de brevets comme variable "proxy" mesurant les flux de connaissance (Jaffe et al., 1998) s'accordent sur le fait que les externalités de connaissance se diffusent plus largement au niveau local, plusieurs limites sont adressées à cette approche. Ces limites sont entre autres, d'ordre technique et sont relatifs à la construction de la base de données, ce qui remet en cause la robustesse des résultats.

Par ailleurs, les citations ne rendent pas compte des connaissances tacites alors qu'elles peuvent constituer une source importante d'externalités (Boufaden, 2008). En réponse à ces imperfections, une approche d'évaluation alternative basée sur la fonction de production de connaissances est proposée. Cette approche présente l'avantage par rapport aux citations de brevets, de pouvoir mesurer les flux de connaissance lorsqu'elles existent et leurs effets sur l'innovation.

2.3. La Fonction de Production de Connaissances

La littérature sur la diffusion des connaissances et la géographie de l'innovation a longtemps cherché une fonction reflétant la production des connaissances et d'innovations. La fonction de production de connaissances ou d'innovations⁴(FPI) a été utilisée au départ pour estimer la relation entre l'intensité de la recherche et développement et la croissance économique. Elle remonte aux travaux de Zvi Griliches (1979) qui est le pionnier dans l'utilisation de cette fonction. Son objectif consistait à étudier le rôle des externalités de connaissance dans la croissance économique. Il considère que la productivité d'une firme dépend non seulement de ses efforts en R&D mais aussi du niveau du stock des connaissances disponibles dans son environnement. Par la suite, Griliches (1992 ; 1994) trouve que les externalités de R&D sont un facteur déterminant de la croissance endogène. En revanche, il montre que la mise en évidence empirique des externalités est une tâche difficile car elle reste tributaire des données disponibles (Boufadden, 2008).

Par ailleurs, plusieurs études ont repris la fonction de production pour tester l'effet des externalités de connaissance mais cette fois sur la production de l'innovation. Jaffe (1989) a été le précurseur en matière d'utilisation de la fonction de production de l'innovation (FPI) dans cette fin. De plus, les premiers travaux sur la fonction de production de connaissances ont été traditionnellement élaborés à des niveaux microéconomiques (Griliches, 1979 ; Baldwin et Scott, 1987 ; Cohen et Levin, 1989). Ainsi, les entreprises sont exogènes et leur performance, en termes d'innovation, est endogène (Scherer, 1984 et 1991 ; Cohen et Klepper, 1992). En effet, les auteurs sont tenté d'expliquer le changement technologique à l'aide de la FPI. Cette fonction stipule que les entreprises sont à la recherche de nouvelles connaissances afin de les intérioriser et de générer des activités innovantes. Ils ont introduit des variables mesurées sur une même échelle tels que l'investissement en R&D, le capital humain, la main d'oeuvre qualifiée (Cohen et Klepper, 1992). Cette fonction suppose que ces intrants généreront des extrants grâce aux activités d'innovations.

Cependant, d'autres travaux sur la FPI sont apparus afin de modéliser les activités d'innovation au niveau régional. Ils soutiennent Scherer (1984) qui stipule que la FPI est beaucoup plus solide lorsqu'elle est utilisée à des niveaux macroéconomiques (Acs et Audretsch, 1988 et 1990). En effet, ces auteurs ont trouvé que les grandes entreprises, c'est-à-dire comptant 500 employés et plus, étaient à l'origine d'un plus grand nombre de produits innovateurs que les plus petites entreprises. Cette comparaison a été basée

4. Appelée aussi Fonction de Production d'Innovation (FPI). Dans ce travail nous utilisons indifféremment les deux notions.

sur une mesure de l'intensité de l'innovation par entreprise, et ceux dans toutes les industries.

Les résultats montrent que la majorité des investissements en R&D industriel est réalisée par les grandes entreprises alors que les petites et moyennes n'investissent que très peu en R&D (Scherer, 1991). Selon ces auteurs, les petites entreprises auraient beaucoup moins d'intrants que les grandes. Ceci suggère la présence de certaines externalités de connaissance ou autres qui se rajoutent aux entrants de la FPI pour produire de l'innovation. Il est donc nécessaire d'aller au delà du niveau désagrégé afin de mesurer les externalités à l'intérieur d'une région. Un premier ensemble de travaux avait pour objectif la mise en évidence des relations inputs/outputs d'innovations tout en modélisant les externalités de connaissance (§ 2.3.1), un autre avait pour objectif de tenir compte de la dimension spatiale des externalités (§ 2.3.2).

2.3.1. Mise en évidence des relations Input/output et indice de coïncidence géographique

Ce travail a été initié par Jaffe (1989) dans son article "*Real Effects of Academic Research*". L'objectif de cette étude a consisté à modéliser l'étendue des externalités de connaissance générées par les universités États-Uniennes. Dans un premier temps, Jaffe (1989) a étudié l'impact de l'activité de recherche des universités sur la capacité d'innovation des firmes de la même zone (mesurée par les brevets). L'échantillon étudié porte sur 8 années et sur 29 États fédérés. L'auteur observe une relation positive et forte entre l'activité d'innovation et à la fois les dépenses en R&D industrielles, et des universités au niveau d'un État. Afin de rendre compte des interactions plus locales entre universités et industries, il a ensuite introduit, dans un deuxième temps, un indice de coïncidence géographique entre la recherche publique et la recherche privée (correspondant à la corrélation entre le nombre de salariés de la R&D privée dans une zone métropolitaine d'un État et le volume de dépenses publiques de recherche effectuée dans cette même aire géographique). L'indice de coïncidence géographique entre la R&D industrielle et la recherche universitaire au niveau de chaque État i est mesuré comme suit :

$$C_i = \frac{\sum_s U_{is} TP_{is}}{[\sum_s U_{is}^2]^{1/2} [\sum_s TP_{is}^2]^{1/2}}, \quad (2.1)$$

avec U_{is} est le montant des dépenses en R&D des universités. TP_{is} , le nombre moyen des chercheurs par laboratoire multiplié par le nombre des laboratoires de R&D

industriel dans chaque MSA_s ⁵ (JTH, 1993 ; Autant-Bernard, 2001a,b).

Les résultats de Jaffe (1989) montrent l'existence des externalités intra-nationales issues de la recherche universitaire en faveur de l'industrie : elles sont favorisées par la coïncidence géographique entre R&D privée et université mais seulement dans un nombre limité de secteurs : secteur médical, optique et la technologie nucléaire.

Par la suite, Acs et al. (1992) reprennent la même méthode que celle de Jaffe (1989) pour mesurer l'incidence de la R&D universitaire sur l'innovation des firmes. Dans leur article "*Real Effects of Academic Research : Comment*", les auteurs ont recours à un nouvel indicateur de l'innovation (le nombre d'innovations enregistrées par la US Small Business Administration sur la base d'articles de journaux spécialisés). Ils ont confirmé l'impact positif de la recherche publique et de la recherche privée d'un Etat sur l'activité d'innovation du même Etat. Leurs résultats confirment ceux de Jaffe quant à l'impact positif de la recherche publique sur l'activité d'innovation d'un Etat : "*the impact of university spillovers is apparently greater on innovations than on patented inventions*" (Acs et al. 1992) et plaident clairement en faveur des effets bénéfiques de la coïncidence géographique (contrairement à Jaffe, 1989). En utilisant le nombre des innovations comme variable output, l'impact de la coïncidence géographique entre universités et unités de R&D privées semble plus important sur les innovations proprement dites que sur le dépôt des brevets. La prise en compte des innovations à la place des brevets renforce donc les résultats de Jaffe (1989). Les R&D publiques sont plus en faveur du secteur électronique que celui du secteur mécanique. En effet, quand nous raisonnons sur les innovations commercialisées, réellement introduites sur le marché, l'observation se situe beaucoup en aval du processus d'innovation tandis que les brevets reflètent un stade préalable au développement de l'innovation. D'autres études ont utilisé la part des ventes réalisées en produits de moins de cinq ans ou encore le taux de renouvellement des produits (Barlet et al 1998).

Cependant, cette mesure est contaminée par un certain nombre de facteurs. Par exemple, en prenant la part des ventes en innovation nous ne tenons pas compte des fluctuations du prix, de l'inflation et de la compétitivité dynamique. En outre, nous supposons que les prix sont fixes et que l'augmentation de produits innovants et de la connaissance entraînent une valeur marchande plus haute, ce qui n'est pas toujours le cas. Cette hypothèse reste donc injustifiée.

Par la suite, et en prenant en compte "*l'infrastructure technologique*", Feldman (1994a) parvient à étendre l'analyse de Jaffe (1989) : aux retombées technologiques issues de la recherche publique sont associées des retombées issues des activités privées

5. En désignant les régions par NUTS 1, 2 ou 3 au niveau Européen, les Etats, les MSAs et les comtés aux USA. Nuts est l'acronyme de la nomenclature d'unités territoriales statistiques de l'Union Européenne.

de recherche mais aussi de production (de service notamment). Elle a trouvé que chacun des inputs a un impact positif et significatif sur l'innovation. Son étude montre également que les *spillovers* sont intersectoriels (résultat trouvé par Jaffe, 1996) et que les petites entreprises bénéficient davantage de la recherche publique et des services locaux que les grandes entreprises. Néanmoins, La dimension locale est simplement mesurée par la localisation à l'intérieur d'un même Etat : elle a considéré les différentes sources de *spillovers* contenues dans ce même Etat. Or, Jaffe (1996) indique que l'Etat n'est pas une échelle géographique adéquate pour rendre compte des *spillovers*.

D'autres analyses (Antonelli, 1994 ; Zucker et al., 1994) se penchent sur les effets de la proximité géographique sur la diffusion des connaissances entre firmes. Leurs résultats confirment l'importance de la proximité géographique, avec d'autres entités en R&D privées dans l'innovation. Par exemple, Antonelli (1994) relie le taux de croissance de la production de 92 firmes italiennes au taux de croissance du stock de capital, de travail et des dépenses de R&D réalisées au sein des firmes et par les autres firmes de la région. Il a trouvé que les activités de R&D privées réalisées dans la même zone géographique ont un impact positif mais faible sur la croissance de la productivité, témoignant d'externalités de connaissance intra-régionales.

Zucker et al. (1994) et Anselin et al. (1997) adoptent une méthode alternative qui consiste à observer si les dépenses de R&D des firmes ou des universités voisines à une firme donnée, affectent la capacité d'innovation de la même firme. Pour cela, ils ont mesuré l'élasticité de l'innovation d'une firme (Ou de groupe de firmes) par rapport aux dépenses de R&D des autres firmes ou des universités situées dans la même zone géographique de référence (ou dans une zone voisine). Cette approche offre une véritable mesure des *spillovers* : l'output d'innovation d'une firme dépend de l'effort de R&D des firmes ou des universités voisines. Elle présente cependant l'inconvénient majeur de postuler le rôle du local.

2.3.2. Prise en compte de la dimension spatiale des externalités de connaissance : le modèle fondateur d'Anselin, Varga et Acs (1997)

Anselin et al. (1997), sont les premiers à perfectionner l'indice de coïncidence géographique calculé par Jaffe (1989). Alors que ce dernier a déduit l'existence de *spillovers* au niveau local en observant l'impact de la recherche menée au sein d'un Etat sur la capacité à innover de ce même Etat, Anselin et al (1997) dépassent le simple effet d'agglomération. Ils ont évalué plus directement l'impact de l'intensité de la recherche menée en dehors de la zone métropolitaine étudiée sur sa capacité à

innover : c'est donc le passage d'une approche qui déduit l'existence des externalités vers une approche qui les mesure de manière plus directe.

Comme nous l'avons mentionné précédemment, Anselin et al (1997) reprennent la méthode de Jaffe (1989), mais en raisonnant au niveau des aires métropolitaines et non plus au niveau des États. Ce faisant, les auteurs considèrent que l'indice de coïncidence géographique n'a lieu d'être utilisé que lorsque l'on mesure les *spillovers* à un niveau géographique trop large. A une échelle d'observation plus petite, il serait possible de mesurer plus directement les *spillovers*.

Le modèle utilisé par Anselin et al (1997) repose sur une fonction de production dans laquelle apparaissent les dépenses de recherches privées et publiques effectuées à l'intérieur et à la périphérie de l'aire métropolitaine. Ils testent trois équations :

$$K = \alpha + R + R75 + U + U50 + LQ + BUS + LARGE + RANK \quad (2.2)$$

$$R = \beta + R50 + U + U75 + HTEMP + FORTUNE + RANK \quad (2.3)$$

$$U = \gamma + R + R50 + U75 + ENRL + EDEXP + RANK \quad (2.4)$$

Avec K, le nombre d'innovations d'une aire métropolitaine donnée, R les effectifs privés de recherche et U les dépenses publiques de recherche. *R75* (respectivement *R50*) mesure les effectifs privés de recherche dans un périmètre de 75 miles (respectivement 50 miles) autour de l'aire métropolitaine et dans un périmètre de 75 miles (respectivement, 50 miles). *LQ* est le nombre d'emplois de haute technologie, *BUS* est le nombre d'emplois dans les services aux entreprises. La variable *LARGE* évalue l'effet de la taille des entreprises. Ensuite, un indicateur de qualité de l'université (*RANK*) est introduit. La variable *HTEMP* constitue un indicateur du degré de spécialisation de l'aire métropolitaine dans les hautes technologies. La variable *FORTUNE* rend compte de la localisation de la R&D à proximité des maisons mères. Enfin, dans l'équation (2.4), *ENRL* et *EDEXP* représentent la taille des universités (nombre d'étudiants inscrits d'une part et les dépenses d'éducation d'autre part).

Ces auteurs, à l'instar de Jaffe (1989) étudient les relations croisées entre la recherche privée et la recherche publique, en faisant dépendre R (recherche privée correspondant au nombre d'employés dans les laboratoires de R&D) de U (recherche universitaire) et inversement. Trois mesures alternatives des interactions spatiales localisées sont appliquées au cas des secteurs de haute technologie :

2.3.2.1. Une mesure de co-localisation

Elle consiste à utiliser le même indicateur de coïncidence géographique que celui de Jaffe (1989) multiplié par le log de la recherche universitaire au niveau de l'Etat, avec une légère modification au niveau du coefficient de corrélation entre R (recherche privée) et U (recherche universitaire). Si Jaffe (1989) a utilisé un coefficient non centré entre R et U pour les zones métropolitaines (MSAs) d'un Etat donné, ces auteurs remplacent le coefficient par la proportion des comtés où la recherche publique est co-localisée avec la recherche privée. Ainsi, pour un Etat s , l'indicateur se calcule de la manière suivante :

$$CUR_s = \frac{\sum \delta R_i \times \delta U_i}{\sum \delta U_i}, \quad (2.5)$$

où δR_i et δU_i prennent respectivement la valeur 1 lorsque $R_i > 0$ et $U_i > 0$ et 0 sinon.

2.3.2.2. Une mesure de gravité

Les auteurs mesurent le potentiel issu de l'interaction entre R&D industrielle d'un comté dans un État donné avec toutes les universités des autres comtés. Cette méthode considère que les interactions économiques et sociales entre deux entités économiques, géographiquement définies, sont proportionnelles à la taille de ces entités et inversement reliées à la distance qui les séparent (Combes et Duranton, 2006).

$$GRAV_i = \frac{\sum_j U_j}{d_{ij}^2}, \quad (2.6)$$

avec i désigne le comté où est localisée la R&D privée au sein d'un État s , et j désigne le comté où est réalisée la recherche publique.

d_{ij} est la distance entre deux comtés. Lorsque la recherche publique et la recherche privée coïncident géographiquement, la variable U_j n'est pas pondérée par la distance d_{ij} .

$GRAV_s$: est la mesure de gravité pour un Etat donné. Elle correspond à la moyenne des $GRAV_i$, lorsque $GRAV_i \neq 0$.

2.3.2.3. Une mesure de couverture

Cette mesure prend en compte les effets sur la R&D industrielle d'un comté, de la recherche publique réalisée dans un périmètre de 50 ou de 75 miles autour de ce comté. Ce qui permet d'introduire un stock de recherche externe, réalisé à l'extérieur de la zone considérée. Plus la distance entre les comtés i et j est faible par rapport à la

distance moyenne de l'ensemble des comtés de l'Etat, plus la coïncidence géographique est forte, et plus l'output d'innovation de l'Etat considéré devrait être élevé.

$$COVi = \sum_j \delta_{ij} U_j \quad (2.7)$$

δ_{ij} prend la valeur 1 lorsque d_{ij} est inférieure à la distance critique de 50 miles ou de 75 miles.

Ces trois mesures utilisées permettent aux auteurs d'établir le constat suivant : leurs premiers résultats (équation 1) confirment ceux de Jaffe (1989). En effet, la recherche universitaire a un impact positif et très significatif sur l'innovation au niveau des zones métropolitaines. Cet impact est beaucoup plus important que celui de la recherche privée.

L'équation (2) indique un effet important de la recherche publique sur la recherche privée, mais seulement à l'intérieur de l'aire métropolitaine. En revanche, des *spillovers* de recherche privée se produiraient entre les aires métropolitaines. Dans l'équation (3) les auteurs mettent en évidence le rôle de la qualité du département de l'université et de la taille de l'université. Les effets de débordement sont plus importants quand l'université est grande et le département de bonne qualité.

Les auteurs affirment que le raisonnement au niveau des comtés au lieu des aires métropolitaines permet de s'affranchir de la mesure de la coïncidence géographique. Néanmoins, ces variables locales U50 et U75 sont en fait des indices de couverture, identiques à celui présenté comme substitut à l'indice de concentration géographique de Jaffe (1989). En plus, en l'absence de mesure de la coïncidence géographique, la méthode de Jaffe ne permet pas de démontrer l'existence de *spillovers* géographiques. Elle permet seulement de confirmer ou non l'existence des *spillovers* et ne précise pas leur dimension spatiale. La réduction de la zone géographique étudiée ne suffit donc pas pour rendre compte des *spillovers* géographiques.

Toutefois, contrairement à l'approche avec coïncidence géographique la méthode présentée par Anselin et al.(1997) présente l'avantage de modéliser les externalités plutôt que de les déduire de la concentration de l'output d'innovation comme pour le cas de Jaffe (1989).

2.3.3. La dimension technologique des externalités de connaissance

L'impact décroissant de la proximité géographique dans le processus d'innovation est donc lié au caractère local des externalités de connaissance. Néanmoins, la prise en compte de la proximité technologique entre firmes, universités, industries, régions ou pays, etc. est fondamentale pour ne pas attribuer la diffusion des externalités à la simple proximité géographique.

Selon Griliches (1992), l'origine de ces externalités c'est le gain lié à la spécialisation (développement du savoir faire et de l'interaction entre artisans et ingénieurs). Glaeser et al. (1992), en étudiant l'effet des *spillovers* technologiques sur la croissance des villes, observent que les industries qui croissent plus dans une ville sont celles qui sont relativement sous représentées, par rapport à la moyenne nationale. Ces externalités utilisées pour analyser la croissance sont aussi appelées externalités dynamiques. Les résultats tendent à confirmer le rôle de la diversification ou externalités de type Jacobs (comme nous les avons défini dans le chapitre 1). Au contraire, les externalités de type MAR ont un effet négatif sur le processus de croissance.

Cependant, pour Glaeser et al., (1992), ce résultat peut être fortement lié au fait que les industries considérées sont des industries à maturité. Selon eux, s'ils avaient pris en compte des industries jeunes, sans doute les externalités intra-industries joueraient-elles davantage. De plus, ces résultats sont contradictoires avec ceux trouvés par Henderson et al., (1995). Ces derniers, dans une étude sur la croissance de l'emploi industriel dans les villes américaines, constatent le rôle positif de la spécialisation (externalités MAR), tandis que la diversité (externalités de type Jacobs) n'a pas d'impact significatif.

L'étude de Paci et Usai (2000) confirme, dans le cas Italien, l'importance des externalités de type Jacobs notamment lorsque la zone dispose d'activités diversifiées de haute technologie. Enfin, les travaux de la géographie de l'innovation s'interrogent plus spécifiquement sur l'origine des externalités locales de recherche. Jaffe (1989) par exemple précise que l'influence de la recherche universitaire est plus nette au niveau des secteurs qu'au niveau local. De même Jaffe et al., (1993), Verspagen et Schoenmakers (2000) montrent que la probabilité de citer un brevet est positivement influencée par la proximité technologique entre le nouveau brevet et le brevet cité.

2.3.4. Quelques prolongements

L'analyse des travaux existants fait ressortir un certain nombre de problèmes liés à la modélisation des externalités et de leur dimension géographique. La mesure des externalités de connaissance pose un certain nombre de problèmes dans la mesure où

il importe de tester deux hypothèses principales :

Hypothèse 1 : Les externalités de connaissance existent.

Hypothèse 2 : les externalités de connaissance sont localisées.

En effet, si les externalités de connaissance ont une dimension locale, ceci doit conduire à la concentration des activités innovantes : les analyses portant sur la concentration en utilisant l'indice de coïncidence géographique reposent sur cette idée. En revanche, s'il y a concentration géographique de l'innovation, cela ne provient pas nécessairement de la présence d'externalités technologiques locales. Ainsi, nous ne pourrions pas conclure avec certitude à la présence d'externalités en confirmant seulement la concentration de l'innovation dans la mesure où cette méthode ne permet que d'observer les conséquences de la concentration et non pas leur présence. En plus, d'autres facteurs, tels que la main d'oeuvre qualifiée peut être à l'origine de la polarisation géographique des activités innovantes.

Nous concluons que les méthodes intégrant un indice de coïncidence géographique sont sans doute les plus à même de traduire le caractère local des externalités. Néanmoins, les externalités ne sont pas modélisées dans ce cas. A l'inverse, les études modélisant les externalités de connaissance traitent pauvrement la dimension géographique. Il paraît dès lors souhaitable d'utiliser une autre méthode qui permet de tester à la fois l'existence et le caractère local des externalités de connaissance.

Les travaux réalisés sur les mesures des externalités de connaissance sont variés et permettent de dégager le rôle de la proximité géographique dans la captation des *spillovers*. Dans l'ensemble, un consensus se dégage autour de l'idée d'une dimension locale des externalités de connaissance. Cependant cette revue de la littérature met en même temps en lumière un certain nombre de difficultés liées à la modélisation des externalités d'une part et de leur dimension géographique d'autre part.

Depuis le modèle d'Anselin et al. (1997), plusieurs modèles basés sur l'interaction spatiale des dépenses en R&D publique versus privée ont vu le jour. Ces modèles tentent de mesurer les externalités tout en tenant compte de leur caractère géographique. Dès lors, les méthodes utilisées pour mesurer les externalités consistent à pondérer le stock de R&D externe par un coefficient qui tient compte de la distance ou de l'ordre de contiguïté entre l'unité considérée et l'unité proche.

2.3.4.1. Nombre de frontières communes entre unités géographiques

Dans une étude sur la France, Autant-Bernard (2001a,b) a testé l'impact des caractéristiques (R&D privée et publique, capital humain) des départements limitrophes d'une part, et des départements voisins de ces mêmes départements limitrophes d'autre part, sur le niveau d'innovation d'un département français.

L'effet de la proximité géographique dans l'innovation dépend de l'ordre de contiguïté entre les départements étudiés. Elle est d'ordre 1 lorsque les deux départements partagent la même frontière ou bien d'ordre 3 lorsque deux départements sont séparés par deux frontières, c'est-à-dire par un troisième département.

Autant-Bernard (2001a,b) trouve que les externalités de connaissance sont géographiquement contraintes dans la mesure où les départements français ne bénéficient que des effets positifs des activités de recherche de leurs voisins proches, et non des activités de recherche plus éloignées géographiquement. Aussi, la diffusion des externalités provenant de la sphère publique est plus sensible à la distance géographique lorsqu'il s'agit d'innover que lorsqu'il s'agit de faire de la recherche : plus on avance dans le processus d'innovation plus les interactions sont importantes et la proximité géographique aussi. D'où la dimension locale des externalités de connaissance.

Enfin, elle trouve que la recherche privée et publique ne se renforcent pas mutuellement. La R&D privée locale n'a pas d'impact sur la recherche publique. D'où la particularité du système français qui est caractérisé par un décloisonnement entre recherche privée et publique ; la R&D universitaire est un input de la recherche privée mais l'inverse n'est pas vrai. Néanmoins, ceci ne veut pas dire pour autant que cette diffusion est contrainte par la distance. En fait, cet impact positif sur l'innovation peut traduire la relation simple et évidente entre les dépenses en R&D internes à chaque département d'une part et les brevets déposés au niveau de chacun de ces départements d'autre part.

Dans un autre travail, Autant-Bernard (2001b) réexamine le problème de mesure des externalités technologiques et de leur diffusion. Elle a montré que le capital humain des départements contigus a un impact sur l'innovation du département considéré. Cet impact est cependant moins important que celui induit par le capital humain local. Mais il est considéré comme étant un vecteur de diffusion des connaissances au niveau local.

2.3.4.2. Découpage par zone géographique

Dans un travail plus récent, Bottazzi et Peri (2003) étudient l'effet de la recherche menée à différents niveaux géographiques : dans la région, dans les régions situées à moins de 300km, dans les régions situées entre 300 et 600km, 600-900, 900-1300, 1300 -2000, sur le processus d'innovation dans de 86 régions européennes. Une variable muette (country) désignant le pays est ajoutée.

Couvrant la période 1977-1995, ces auteurs constatent que les externalités de R&D sont localisées et de très faible intensité. Ils trouvent que seule la R&D des régions situées à une distance inférieure à 300 km a un impact sur la capacité à innover d'une

région donnée : lorsque les dépenses en R&D d'une région doublent par exemple, le nombre de brevets déposés des autres régions situées dans cet intervalle va augmenter de 2-3%, alors que cela augmente de 80-90% les innovations de cette même région.

Par la suite, ces auteurs prennent en compte la distance d'une manière paramétrique. En attribuant un poids à la R&D sur différentes distances par rapport à la région considérée, les auteurs construisent une variable "pool" de la R&D. Cette variable représente la contribution effective de chaque région à l'innovation de la région i compte tenu de la distance qui les sépare. La moyenne de chaque variable $lg(R\&D)$ pour chaque 100km est ensuite pondérée par une fonction distance $F(d)$. Cette fonction est spécifiée de trois manières différents : linéaire ($F(d) = 1 - (d/20)$), inverse ($F(d) = (1 + d)^{-1}$) et exponentielle ($F(d) = e^{-d}$). Les auteurs trouvent qu'une augmentation de 1% dans la R&D génère un effet externe de 0.07% sur la capacité à innover d'une région à une distance de 0 km. Les externalités diminuent très rapidement telle que à 200 km, l'impact est de 0.008% et nul à 500km.

Par ailleurs, si certains travaux ont contrôlé seulement la proximité géographique (Bottazzi et Peri, 2003 ; Maïresse et Mulkay, 2007), d'autres ont contrôlé aussi la proximité technologique (El Ouardighi, 2002 ; Abdelmoula et Bresson, 2007 ; Boufaden et Plunket, 2008). Ces derniers travaux ont montré que les effets de *spillovers* géographiques diminuent avec la prise en compte de la proximité technologique.

Ainsi, Abdelmoula et Bresson (2007) utilisent une mesure de proximité technologique proche de celle de Jaffe (1986). Ils exploitent la distribution des brevets d'une région dans les différentes classes technologiques pour caractériser la position de cette région dans les différentes classes technologiques. Cette matrice est une moyenne des proximités obtenue à l'aide des demandes de brevets de 1995 à 2002 désagrégées sur les sections internationales de la classification des brevets (International Patent Classification(IPC)). Ces auteurs trouvent, contrairement aux autres études sur les *spillovers*, que pour les deux matrices utilisées (géographiques et technologiques), les effets de *spillovers* sont significatifs et peuvent être négatifs selon les distances entre les régions européennes. Au final, il semble que les dernières approches sur la FPI sont plus précises et permettent de mieux appréhender le caractère local des externalités de connaissance.

Même si nous admettons l'importance de ces méthodes dans la mesure des effets d'externalités dans l'innovation, il demeure difficile d'admettre ce lien positif entre externalités et innovation. En effet, la plupart des études montrent un effet positif des externalités sur l'innovation sans leur donner une définition stricte. Elles ne font pas une distinction entre externalités de connaissance et externalités de R&D à l'exception de l'étude faite par Abdelmoula et Bresson (2007). Ces travaux utilisent les dépenses

en R&D pour mesurer des externalités et les assimilent généralement à des externalités de connaissance.

Comme nous l'avons montré dans le chapitre 1, les externalités de R&D peuvent constituer un frein à l'innovation dans la mesure où elles posent le problème de l'appropriation des connaissances (Autant-Bernard, 2000). Même si ces externalités sont positives, nous ne pourrions pas attribuer l'effet positif des externalités aux externalités de connaissance au sens strict. En effet, en mesurant les externalités de connaissance, il est difficile de faire abstraction des externalités pécuniaires (Cheynet et Fadaïro, 1998). En même temps, la R&D n'est pas la seule origine des *spillovers* puisque une partie de ces *spillovers* ne se diffuse qu'une fois que l'innovation a vu le jour, et non pas pendant la phase amont de la recherche.

Pour toutes ces raisons, nous ne pouvons dire que les travaux sur les *spillovers*, même les plus récents, mesurent adéquatement les externalités de connaissance. En effet, même si ces travaux sont de plus en plus développés et que la distinction entre proximité géographique et technologique employée a permis de relativiser l'effet des externalités géographiques au profit des externalités technologiques, nous ne pouvons pas savoir s'il s'agit vraiment d'externalités de connaissance, pécuniaires ou plutôt de R&D.

Pour toutes ces raisons, nous mesurons dans ce qui suit les "*flux de connaissance*" dans des intervalles de distances différents. Ces flux peuvent être des externalités de connaissance, des externalités de R&D, des externalités technologiques, etc.

Contrairement à ce qui a été évoqué dans les études antérieures, nous considérons que les externalités géographiques peuvent être technologiques en même temps. Ainsi, comme nous l'avons mentionné dans le chapitre 1, notre objectif principal ne consiste pas à distinguer les externalités géographiques des externalités technologiques car il nous semble que les méthodes adéquates pour ce faire ne sont pas assez développées. Nous ambitionnons plutôt de montrer l'effet de la proximité géographique dans le transfert des connaissances à travers l'introduction des variables réseaux. Nous supposons que la proximité géographique n'est favorable aux transferts des connaissances qu'à travers la prise en compte des canaux de diffusion de ces connaissances.

2.4. Cadre d'analyse empirique et données

Nous présentons dans cette section les hypothèses construites⁶ dans notre partie empirique ainsi que les différentes bases de données exploitées.

2.4.1. Hypothèses et description de l'étude

Comme mentionné dans le chapitre premier, les flux de connaissance ne sont ni entièrement localisés, ni complètement libres ou diffusables à n'importe quelle distance et entre tout type d'agents ou lieux. Partant de ce constat, l'effet des connaissances mesurées dans les différents cercles concentriques dépendra nécessairement de trois facteurs : 1) le type de la connaissance en jeu ; 2) le type du canal de diffusion qu'elle utilise et 3) les caractéristiques des agents, des organisations ou des lieux impliqués dans le processus d'échange des connaissances.

En ce qui concerne le premier facteur, nous avons montré précédemment que la distinction entre connaissance tacite versus codifié a souvent été utilisée comme argument expliquant pourquoi les *spillovers* semblent délimités dans l'espace, mais aucune étude n'a comparé l'étendue spatiale des connaissances tacites à celle des connaissances codifiées.

Dans la littérature empirique sur les *spillovers*, plusieurs études proposent de distinguer la R&D privée de la R&D publique ou bien la R&D académique de la R&D entrepreneuriale. Ces distinctions ne semblent pas très utiles pour comprendre pourquoi certains flux de connaissance sont localisés alors que d'autres peuvent se diffuser à des distances plus longues. Dans la théorie, la distinction entre connaissances codifiées et connaissances tacites, nécessitant plus de contacts face à face, nous permettrait ainsi de comprendre pourquoi certaines connaissances sont plus étendues dans l'espace que d'autres.

Pour cette raison, nous construisons deux mesures des dépenses en R&D. Nous proposons une approximation raisonnable consistant à mesurer les connaissances codifiées par la R&D externe et les connaissances tacites par la R&D interne. En effet, les entreprises qui cherchent à tirer profit des connaissances tacites investissent en R&D interne : "*The benefits of such tacit knowledge arise only through a culture of trust and knowledge-sharing within an organization*" (Cowan et al., 2000, p. 223). Au contraire, la R&D externe s'appuie très fortement sur les connaissances codifiées qui sont plus

6. Nous reprenons ici les arguments, hypothèses et propositions développés dans Brossard et Moussa (2012), *Papers in Evolutionary Economic Geography, May 2012*.[http : //econ.geo.uu.nl/peeg/peeg.html](http://econ.geo.uu.nl/peeg/peeg.html).

facilement transférables que les connaissances tacites. Notre première hypothèse est donc la suivante :

H1 : La R&D interne intègre à la fois des connaissances tacites et des connaissances codifiées alors que la R&D externe ne véhicule pratiquement que des connaissances codifiées. Par conséquent, en introduisant ces deux mesures différentes dans notre modèle, nous pourrions distinguer les limites géographiques de chacune des connaissances tacites et codifiées.

Proposition 1 : La nature des connaissances a un impact significatif sur l'ampleur des externalités de connaissance. Par conséquent, la R&D interne devrait avoir un effet plus élevé sur l'innovation que la R&D externe, puisque cette dernière exploite essentiellement des connaissances codifiées tandis que la première intègre aussi des connaissances tacites.

La seconde série de facteurs que nous souhaitons intégrer dans l'analyse concerne les canaux de transmission des connaissances. Nous avons montré dans le chapitre premier qu'il existe trois canaux de diffusion des connaissances à savoir la mobilité des travailleurs (entre entreprises), les spin-off (entre une filiale mère et de nouvelles entreprises) et les réseaux sociaux (transfert des connaissances dans des réseaux sociaux et à un niveau plus agrégé). Nous nous concentrons sur ce dernier canal puisque les réseaux sociaux offrent de nombreux canaux de diffusion des connaissances formelles et informelles au niveau des régions (Tappeiner et al, 2008). Ainsi, tout type de réseau permet une interaction face à face ou bien à distance entre chercheurs hautement qualifiés ou bien entre ingénieurs.

H2 : les clusters dotés des réseaux de diffusion des connaissances les plus variés et les plus efficaces ont une meilleure productivité à l'innovation car ils tirent partie des spillovers de connaissance plus intenses émanant de ces réseaux. Dans les clusters high-tech, ces différents canaux sont complémentaires et ont des effets cumulatifs plutôt que substituables.

Proposition 2 : Les différents canaux de diffusion de la connaissance ont des effets cumulatifs sur la production de l'innovation. En conséquence, les clusters dotés d'un grand nombre de réseaux de bonne qualité auront une meilleure productivité à l'innovation.

En d'autres termes, certains départements ont plus de clusters que d'autres et ces clusters sont mieux dotés en canaux de transmission de la connaissance. Ceci aura bien évidemment un effet sur la production d'innovation de chaque département. En France, les pôles mondiaux et à vocation mondiale sont des clusters high-tech particulièrement

bien dotés en réseaux dans la mesure où ils peuvent attirer des ingénieurs et des chercheurs hautement qualifiés.

La troisième série de facteurs concerne le débat spécialisation/diversification. Certaines études indiquent que les flux de connaissance se produisent principalement entre les entreprises d'une même industrie (externalités de type MAR). Au contraire, d'autres indiquent que la diversité technologique est plus favorable à la production des externalités de connaissance (externalités de type Jacobs). Beaudry et Schiffauerova (2009) montrent que le niveau de désagrégation géographique influe fortement sur les résultats, ce qui suggère que la spécialisation et la diversification peuvent influencer la diffusion des connaissances à différentes distances. Les nouvelles connaissances peuvent être difficiles à trouver dans un environnement local hautement spécialisé. Par conséquent, les agents économiques peuvent acquérir de nouvelles connaissances à travers des interactions avec des agents plus distants. Ainsi, si la diversité n'est pas assez importante dans une région, celle-ci peut bénéficier de la diversité de ses voisins.

H3 : Puisque les industries similaires ont tendance à choisir des localisations similaires, la recherche des connaissances dans d'autres secteurs peut nécessiter des interactions à distance avec les détenteurs des connaissances situés dans des régions voisines. Il importe donc pour l'innovation régionale que les régions voisines soient diversifiées plutôt que hautement spécialisées.

Proposition 3 : la diversité départementale a un effet positif sur le département lui-même et aussi sur ses voisins. Les régions diversifiées offrent de la nouveauté exploitable pour l'innovation à la fois à elles-mêmes et aux régions voisines.

Afin de tester ces hypothèses, nous proposons d'utiliser l'approche empirique la plus connue pour étudier les externalités de R&D. Comme mentionné dans le chapitre premier, cette méthode a été formalisée par Jaffe (1989). Ses résultats montrent que les externalités sont localisées et qu'elles présentent une grande sensibilité à la proximité géographique. Plus tard, Bottazzi et Peri (2003) ont appliqué cette approche sur d'autres données pour montrer que les externalités de R &D sont localisées et de très faible intensité.

En ce qui nous concerne, nous partons de cette dernière étude basée sur le découpage géographique. Nous estimons une fonction de production d'innovation (FPI) et nous essayons de mesurer l'effet des flux de connaissance externes d'un département français donné sur sa production d'innovation. Nous choisissons au début des classes de distances géographiques d'ampleur 200 kilomètres comme le montre la

figure (2.2) suivante :



Figure 2.2 – Mesure des externalités dans des intervalles de 200 kilomètres

Ensuite, et afin de mieux borner les effets des flux de connaissance externes, nous utiliserons plus de classes de distance, avec des intervalles d'amplitude moindre ($]0 - 100]$, $]100 - 200]$, ...).

Notre spécification est donc une extension de celles de Bottazzi et Peri (2003). Nous adoptons dans un premier temps la même méthodologie de cercle concentrique appliquée sur des données en coupe (§2.5). Ensuite, cette même approche est appliquée sur des données de panel (§2.6). L'objectif de cette dernière application est d'une part, de tester la robustesse de la méthode de cercle concentrique utilisée par certains travaux sur les externalités de connaissance et d'autre part, de tenir compte de l'hétérogénéité des observations.

L'objectif principal de ce modèle est de permettre aux connaissances incorporées dans la R&D de la région j d'avoir des impacts différents sur l'innovation de la région i selon la distance géographique entre i et j .

Nous analysons une coupe de 94 départements français métropolitains (régions NUTS 3). Nous faisons abstraction de la Corse pour laquelle la notion de distance géographique est délicate et ne peut être abordée par les méthodes utilisées dans notre travail. La distance géographique entre les départements est représentée par la distance en kilomètre entre les deux centroïdes de chaque département⁷. L'unité de mesure pour notre cas est la distance à vol d'oiseau⁸. La distance entre les départements varie de 7,94 km (entre Paris (75) et Hauts-de-Seine (93)) à 1004,771 km (entre le Finistère

7. Cette distance peut être mesurée en utilisant les capitales pour le cas des pays ou bien dans le cas d'utilisation des observations réduites à un point. Ce point est développé en détail dans la section matrice des poids du chapitre 4.

8. L'unité de mesure servant à définir la distance peut être aussi la distance parcourue par route, le temps de transport (heures de vol par exemple ou heure de voiture ou de train).

(29) et les Alpes Maritimes (06)).

Les départements varient en taille. Le plus petit est Paris (75) avec une superficie de 105 km^2 et le plus grand est la Gironde avec une superficie de $10\,000 \text{ km}^2$. Si nous supposons abusivement que ces aires sont circulaires, la plupart des départements ont alors un rayon qui varie entre $5,783 \text{ km}$ et $56,433 \text{ km}$.

Nous donnons dans ce qui suit les statistiques descriptives sur ces intervalles. Bottazzi et Peri (2003) ont choisi cinq classes de distances géographiques et supposent une absence de diffusion au delà de 2000 km . Comme nous travaillons sur un niveau géographique plus petit (NUTS 3), nous avons choisi de diminuer les intervalles de sorte que chaque département ait au moins un département dans chaque intervalle. Nous mesurons les connaissances dans des classes de 200 kilomètres puis dans des classes géographiques de 100 kilomètres. Notre choix implique une absence de diffusion au delà de 600 kilomètres⁹ :

]0 – 200] nombre moyen de départements : 16,298
]200 – 400] nombre moyen de départements : 33,085
]400 – 600] nombre moyen de départements : 29,425
]plusque600km[nombre moyen de départements 14,213

Dans un deuxième temps, nous analysons les effets de la R&D dans des intervalles plus petits afin de mieux les borner :

]0 – 100] nombre moyen de départements : 4,638
]100 – 200] nombre moyen de départements : 11,660
]200 – 300] nombre moyen de départements : 15,468
]300 – 400] nombre moyen de départements : 17,617
]plusque400km[nombre moyen de départements : 43,638

Après avoir donné un aperçu du cadre de notre analyse, nous présentons dans la sous section suivante les trois sources principales de nos données.

2.4.2. Données

Nos variables ont été construites à travers l'utilisation de trois types de données. Les données de l'enquête R&D, les données de l'enquête sur les pôles de compétitivité réalisée par la Direction Générale de la Compétitivité de l'Industrie et des Services (DGCIS) et les données de brevets de l'Institut National de la Propriété Intellectuelle (INPI).

9. Nous avons estimé les mêmes spécifications en considérant que la dernière classe englobe les départements distants de plus de 600 kilomètres. Ceci ne change pas nos résultats.

2.4.2.1. La base de données R&D

Les dépenses en recherches et développement sont issues de l'enquête R&D du ministère de la Recherche¹⁰.

L'enquête annuelle R&D est une enquête auprès des entreprises qui, potentiellement, font de la recherche. Elle recense les entreprises ayant au moins un chercheur en équivalent temps plein de recherche (etp). Les entreprises avec moins d'un chercheur en (etp) ne sont pas prises en compte. Néanmoins, différents forçages automatiques ont été réalisés dans l'enquête. Par exemple, si une entreprise déclarait des effectifs de recherche mais pas de chercheurs, un forçage était automatiquement effectué : dès qu'il y avait au moins 5 personnes en effectif de R&D, l'entreprise se voyait automatiquement attribuer un chercheur.

Ceci évite de compter les entreprises qui font de l'innovation et qui brevètent alors qu'elles ne disposent pas de chercheurs en etp (Même s'il est rare de breveter quand l'entreprise ne dispose pas de chercheurs).

Nous disposons également de l'enquête communautaire sur l'innovation : CIS3 (1998-2000) et CIS 4 (2002-2004). Ces enquêtes auraient pu être utilisées mais elles présentent certaines limites par rapport à l'enquête R&D.

Pourquoi nous utilisons l'enquête R&D au lieu de CIS ?

L'utilisation de l'enquête CIS3 et CIS4 pour modéliser les données de l'innovation est marquée par une limite principale : la non représentativité au niveau départemental. Autrement dit, le nombre des entreprises enquêtées dans chaque département n'est pas proportionnel au nombre total d'entreprises existantes dans ce même département. D'une manière générale, pour que chaque observation soit représentative dans une enquête, nous devons la pondérer par un coefficient de pondération. Ce coefficient permet d'obtenir des estimations non biaisées : c'est ce que nous appelons redressement ou pondération d'une enquête.

Plusieurs méthodes de redressement sont généralement utilisées pour corriger le biais de sélection.

Une première méthode appelée "*méthode de repondération*", consistait à modifier le poids de chaque observation en calculant un nouveau coefficient égal à : $\frac{1n_i}{NN_i}$
Avec N : le nombre total des entreprises au niveau national; n_i : le nombre des entreprises dans l'échantillon et N_i : le nombre total d'entreprises dans le département i .

10. Autorisation du comité du secret statistique de l'INSEE, décision du 17 juin 2010, point 4.2.4 (a).

Une deuxième méthode appelée méthode du "*Raking Ratio*" a aussi été utilisée. Cette méthode est plus complexe que la précédente et a été employée par l'INSEE jusqu'à présent pour redresser un bon nombre d'enquêtes¹¹.

Les résultats trouvés après correction du biais de sélection des enquêtes CIS auraient pu être utilisés mais nous avons rencontré un deuxième problème aussi important que la représentativité des données : celui de la répartition des dépenses de R&D entre établissements d'une même entreprise. Nous avons donc dispatché les dépenses en R&D en utilisant la part d'effectif salarié de l'établissement considéré, dans l'effectif salarié total du département. Ceci a biaisé nos résultats et les a rendus moins fiables dans la mesure où nous ne disposions pas d'information sur la répartition du nombre de chercheurs par établissement.

Une troisième limite de l'utilisation des données CIS réside dans la fusion de plusieurs enquêtes (CIS3 avec CIS4 et EAE). Par manque d'informations, nous avons utilisé l'enquête EAE pour déterminer les codes postaux de chaque entreprise. Ceci a réduit énormément le nombre d'observations : nous sommes passés de 20672 à 2197 observations.

Tous ces problèmes semblent résolus avec l'enquête R&D dans la mesure où les informations disponibles dans l'enquête annuelle sur les moyens consacrés à la R&D sont réparties dans trois fichiers correspondant à trois niveaux d'agrégation :

- **BASENT** comprend les données agrégées au niveau de l'entreprise. Il contient quelques variables recherches (Dird, Derd) et il ne comprend qu'un seul enregistrement pour chacune de ces entreprises.
- **BRAN** comprend les données agrégées au niveau des branches d'activité de R&D de l'entreprise. Il contient plus d'enregistrements que le précédent mais généralement une entreprise ne fait de la recherche que dans une seule branche. Ce fichier comporte l'ensemble des variables proprement recherche (effectif de R&D, dépenses, financements extérieurs, types de financement extérieur).
- **BREG** comprend les données au niveau le plus fin, c'est-à-dire au niveau départemental de chaque branche d'activité R&D de l'entreprise. Ce fichier contient plus d'enregistrements que le fichier BASENT car une entreprise peut faire de la recherche dans plusieurs régions/départements. Cette ventilation des dépenses de recherche n'est possible que pour les entreprises recevant un questionnaire général. Pour les entreprises recevant un questionnaire simplifié (QS), la recherche est affectée en totalité à un seul département.

11. Dans notre cas, nous avons choisi la première méthode vu que la deuxième est généralement utilisée dans le cas d'échantillonnages complexe ou bien pour des sous-populations.

Ces trois fichiers sont diffusés jusqu'à l'année 2006. A partir de 2007, la diffusion des données sur les moyens consacrés à la R&D se fait par le biais d'un seul fichier et non plus de trois. Ce dernier regroupe des variables relatives aux trois niveaux d'agrégation des données (BASENT, BRAN et BREG).

Pour conclure, l'enquête R&D présente quatre avantages par rapport à l'enquête CIS. Tout d'abord, le problème de répartition de la R&D est résolu si on utilise le fichier BREG. Deuxièmement, les données de l'enquête R&D sont représentatives au niveau sectoriel et au niveau de la taille des entreprises tant à l'échelle nationale que régionale. Troisièmement, les dépenses en R&D externes sont réparties en R&D externe privée et R&D externe publique. Finalement, le code postal de chaque établissement est disponible dans l'enquête R&D ce qui évite le recours à l'appariement avec d'autres enquêtes afin de déterminer la localisation géographique de chaque établissement. Pour toutes ces raisons, nous mobilisons les données de l'enquête R&D pour repérer les ressources nécessaires à l'innovation.

2.4.2.2. La base de données brevets

Les données de toutes les demandes de brevets publiées d'origine française ont été notre mesure empirique de l'innovation d'un département français. Ces données sont fournies par l'Institut National de la Propriété Intellectuelle (INPI). En effet, les données disponibles sur le site de l'INPI sont incomplètes pour notre étude puisqu'elles ne recensent pas toutes les inventions d'origine française. Suite à notre demande d'une base plus complète, l'INPI nous a fourni les données de toutes les demandes des brevets publiées d'origine française déposées par une voie autre que la voie française, notamment les voies européennes et américaines. Ces données sont issues, au moins pour une partie, d'une recherche effectuée sur le territoire français, quelle que soit la nationalité du déposant (entreprise, organisme de recherche, inventeur indépendant, etc.).

Les données de brevets dont nous disposons ont donc l'avantage d'être recensées quelque soit l'office auprès duquel a été effectuée la demande initiale, en évitant de compter deux ou plusieurs fois la même invention¹².

Autrement dit, une demande de brevet déposée auprès de l'INPI, de l'Office Européen des brevets (OEB) ou du Traité de Coopération en matière de brevets

12. Pour éviter le double comptage des brevets, une première analyse est faite et consiste à recenser les demandes de brevet sans priorité déposées à l'INPI, les demandes directes à l'OEB qui ne revendiquent pas de priorité et les demandes sans priorité à l'OMPI, ainsi que les dépôts sans priorité effectués auprès d'un autre office.

(PCT)¹³ ne peut être comptabilisée qu'une seule fois, ce qui évite de compter deux ou plusieurs fois la même invention¹⁴.

De plus, ces données représentent seulement les demandes de brevets publiées, ce qui évite de compter des brevets qui ont été déjà rejetés ou retirés.

En outre, les données collectées de l'INPI présentent une autre particularité par rapport aux données produites par l'Office Européen des Brevets (OEB) ou par l'Organisation Mondiale de la Propriété Intellectuelle (OMPI). Si ces derniers comptabilisent les données selon l'adresse de résidence du déposant, l'INPI utilise celle de l'inventeur¹⁵.

Pour chacune des demandes de brevets publiées, la localisation de l'inventeur nous permet de situer le département à l'origine de l'innovation, ce qui est plus cohérent pour mesurer la localisation de l'innovation. En outre, l'utilisation de l'adresse du déposant aurait pour conséquence de biaiser la localisation en faveur de la région Île-de-France ou se trouvent de très nombreux sièges d'entreprises.

Enfin, si un brevet liste plusieurs inventeurs, la méthode de comptage fractionnel est appliquée afin d'éviter un double comptage. Ainsi, si une demande liste trois inventeurs de trois régions différentes, l'INPI octroi 1/3 à chaque département. Nous obtenons ainsi des décimales dans nos données de comptage des brevets.

Ainsi, nous utilisons une base de données de brevets qui mesure les inventions d'origine française quelque soit l'office de brevets auquel la demande a été déposée.

En tenant compte des brevets d'origine française et déposés auprès des offices autre que l'INPI, le nombre de brevet par département de notre base de données devient plus important que celui qui figure sur le site de l'INPI. Nous présentons une comparaison de statistiques dans les annexes de ce chapitre.

13. C'est un traité international administré par l'OMPI, permet de demander un brevet simultanément dans un grand nombre de pays en déposant une seule demande.

14. Le brevet peut être obtenu par une procédure nationale, qui aboutit à un brevet national, mais aussi par une procédure régionale ou internationale qui aboutit à un ensemble de brevets nationaux. Par exemple la convention de Paris donne aux inventeurs un délai de priorité d'un an, à partir du premier dépôt, pour procéder aux autres dépôts. Ainsi, on court le risque de comptabiliser deux ou plusieurs fois des demandes qui correspondent en fait à la même invention.

15. L'OEB et l'OMPI comptabilisent les demandes de brevets selon la résidence du déposant telle qu'elle apparaît sur la demande : s'agissant d'une personne morale, il s'agit donc du siège social de l'entreprise titulaire du droit à la date de dépôt. Or, il est fréquent que cette résidence soit différente du lieu de l'invention pour plusieurs raisons (organisation, simplicité de gestion, fiscalité, etc.). c'est ainsi, par exemple, que les statistiques usuelles attribuent aux Pays-Bas, à la Finlande ou à la Suisse des demandes de brevets effectuées respectivement par Philips, Nokia ou Nestlé alors que souvent les inventeurs travaillent et résident, non pas dans le pays où se trouve la société mère, mais dans une autre filiale située dans un autre pays.

2.4.2.3. Les données sur les pôles de compétitivité fournies par la Direction Générale de la Compétitivité, de l'Industrie et des Services (DGCIS)

Pour construire nos variables de réseaux, nous exploitons un survey réalisé par la Direction Générale de la Compétitivité, de l'Industrie et des Services (DGCIS). Ce survey sert à évaluer l'impact de la politique des pôles de compétitivité sur le développement des collaborations entre acteurs du processus d'innovation¹⁶. Les données s'appuient sur les bases de données suivantes :

- Les bases extraites de l'enquête sur les pôles de la DGCIS : ces bases fournissent des données de caractérisation des pôles et les données de cadrage sur les projets collaboratifs ayant reçu des financements publics. La liste des établissements n'étant pas connue, d'autres bases sont nécessaires pour étudier les coopérations au sein des projets de R&D des pôles.
- Les bases sur les projets de recherche en partenariat financés dans le cadre des pôles par le Fond Unique Interministériel (FUI)¹⁷. Cette base a été fournie à EUROLIO par la DGCIS.

L'objectif de ces bases est d'analyser l'impact de la politique des pôles sur la structuration des relations entre les différents membres d'un pôle et entre le pôle et l'extérieur. L'intérêt de ces données est de pouvoir mesurer les relations entre les membres selon leur statut (entreprise, laboratoire, autres,...), leurs caractéristiques (taille des entreprises par exemple pour voir l'implication des PME) et leur localisation.

Nous utilisons ce survey de DGCIS-EUROLIO pour construire nos variables réseaux notamment la variable mesurant les co-publications produites par chaque pôle mondial et à vocation mondiale entre 2002 et 2009, ainsi que la variable mesurant la densité des réseaux régionaux dans chacun de ces pôles¹⁸.

16. Nous remercions l'observatoire Eurolio (European Localized Innovation) de nous avoir gracieusement fourni ces données : DGCIS-EUROLIO, octobre 2011.

17. Cette base contient aussi des bases de projets financés dans le cadre de l'Agence Nationale de la Recherche (ANR), dans le cadre de la signature de convention de confidentialité réglementant leur utilisation. Nous nous limitons dans ce travail à l'utilisation des bases financées par le FUI qui semblent suffisantes pour construire nos variables réseaux.

18. Voir la définition de ces variables dans la section (2.6.1).

2.5. Application de l'approche par cercles concentriques sur données en coupe

2.5.1. Variables et modèle

Nous présentons dans cette partie les variables utilisées dans notre étude en coupe ainsi que les modèles. Comme les départements sont différents par tailles géographiques, nous devons standardiser les variables (dépenses en R&D et brevets) par les surfaces en espérant que cela éliminera l'effet taille, cause possible de distorsion (Bottazzi et Peri, 2003; Abdelmoula et Bresson, 2007).

2.5.1.1. La variable dépendante

L'output d'innovation retenu dans notre cas n'est autre que les brevets. Au delà des limites traditionnelles attribuées à cet indicateur, il reste tout de même pertinent pour mesurer l'activité d'innovation. A la fin des années 1970, la disponibilité de l'information en forme lisible sur machine auprès de l'office Américain des brevets, a permis l'utilisation de ces données avec de plus grands échantillons d'entreprises. Griliches (1984) puis Griliches et al. (1987) et Griliches (1990) ont, par la suite, utilisé ces données dans leurs études sur l'innovation. Parallèlement à ces travaux sur les Etats-Unis, d'autres sur l'Europe sont apparus comme par exemple ceux de Schankerman et Packes (1986) qui ont ouvert la voie d'utilisation des brevets de l'OEB.

Cependant, de nombreuses études empiriques ont indiqué que les entreprises (l'industrie pharmaceutique constituant une exception notable), ne considèrent pas le brevet comme un outil efficace pour s'appropriier leurs connaissances. Scherer et al. (1959) aux Etats-Unis et Taylor et Silberston (1973) au Royaume-Uni, confirmés ensuite par Cohen et al. (2000) aux Etats-Unis, au Japon par Goto et Nagata (1996), en Europe par Arundel et Van de Paal (1995) puis Arundel (2001) ont montré que les entreprises de toutes les tailles préfèrent le secret au brevet pour se protéger. Les brevets ne sont pas donc sans limites. En effet, il est essentiel de souligner que toutes les innovations ne sont pas brevetables et que tous les brevets ne débouchent pas sur une innovation¹⁹.

Ensuite, pour observer les innovations dans le temps, nous avons besoin de faire l'hypothèse que la probabilité de breveter est la même tout le temps. Ceci n'est pas toujours vrai. En effet, un changement dans la politique des brevets affecte le taux de "*brevetabilité*" et provoque un grand problème d'estimation.

19. Ce phénomène peut s'expliquer par exemple par l'insuffisante rentabilité de l'invention, qui conduit à abandonner le projet et à renoncer à l'introduction de l'invention sur le marché.

Finalement, une autre difficulté liée à l'utilisation des brevets comme indicateur de l'innovation réside dans les très fortes variations de propension à breveter d'un secteur à l'autre. En effet, pour comparer les données entre industries, les travaux sur les brevets doivent faire l'hypothèse que la probabilité de breveter est la même pour toutes les industries ; cette hypothèse n'est pas vraiment réaliste. Par conséquent, nous pourrions nous attendre à une baisse des dépôts de demandes de brevets. Mais, la réalité est toute autre²⁰.

Malgré ses lacunes, plusieurs études ont été entreprises par les brevets. En effet, les brevets procurent des droits de monopoles aux entreprises en conséquences de leur invention et de leur publication. Ces avantages sont au nombre de trois. Le premier consiste en l'étendue des séries temporelles. En effet, les brevets sont une source d'information facilement disponible et pour de longues séries temporelles. Le second consiste en l'étendue des détails correspondants aux brevets. Le troisième avantage est sans doute le plus important : les inventions doivent répondre à un standard minimum de qualité pour être brevetés. En effet, pour être breveté, une innovation doit être originale, utile et non évidente.

Pour toutes ces raisons, nous exploitons les données de brevets afin de mesurer l'output d'innovation des 94 départements français métropolitains. Ces données représentent la moyenne des années 2006-2007-2008. Le nombre moyen de brevets par département est 186,31 brevets :

20. Nous avons utilisé la base de données de l'OMPI pour étudier l'évolution des demandes de brevets européens au cours de la période 1989-2006. Le diagramme obtenu montre une augmentation des demandes de brevets par année de dépôt dans le monde. Cela se comprend du fait que le brevet permet de diffuser la connaissance, et assure ainsi la coordination des acteurs du processus d'innovation (Bureth et al., 2005)

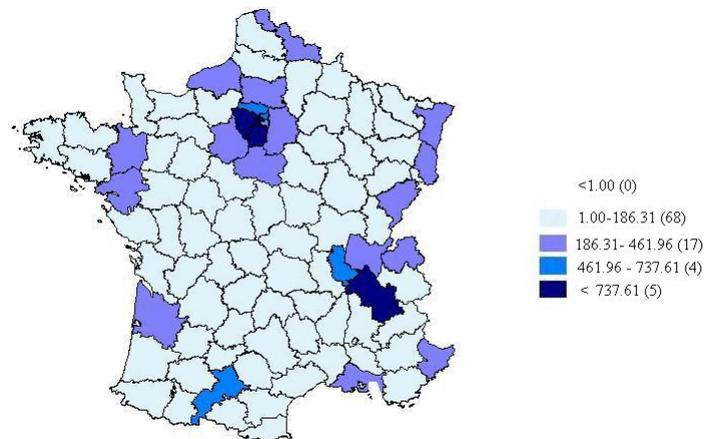


Figure 2.3 – Répartition du nombre moyen de brevets mesuré entre 2006 et 2008

Le nombre de brevets moyen est maximum pour le département Hauts-De-Seine (1452) suivi de Paris (1242). Pour d'autres départements comme le département Lozère, le nombre moyen est 1.

2.5.1.2. Les variables indépendantes

Recherche et développement : Pour mesurer la R&D propre à chaque département, nous avons calculé la R&D départementale interne et la R&D départementale externe. Certaines études ont utilisé seulement la R&D interne exécutée par l'entreprise elle-même afin de mesurer l'effort propre à chaque entreprise ou à chaque zone (Autant-Bernard, 2001a,b; Varga, 1998). Néanmoins, ces dépenses ne représentent pas la totalité de la R&D engagée par l'entreprise ou par la zone étudiée. Par conséquent, une valeur non négligeable et parfois très importante de la R&D n'est pas prise en compte dans l'estimation des modèles qui testent généralement l'effet de la R&D dans le processus d'innovation.

Nos dépenses en R&D sont réparties en trois catégories : R&D interne, R&D externe privée et R&D externe publique.

- La R&D interne (dird)

Nous avons rassemblé la R&D interne établie initialement par établissement en utilisant le code postal correspondant à chaque établissement. Ceci est plus homogène pour mesurer la localisation de la R&D investie dans chaque département. En effet, une entreprise peut avoir plusieurs établissements qui ne sont pas nécessairement localisés dans le même département que le siège de l'entreprise.

En moyenne, le montant d'investissement en R&D interne par département est de 2.351.180 milliers d'euros :

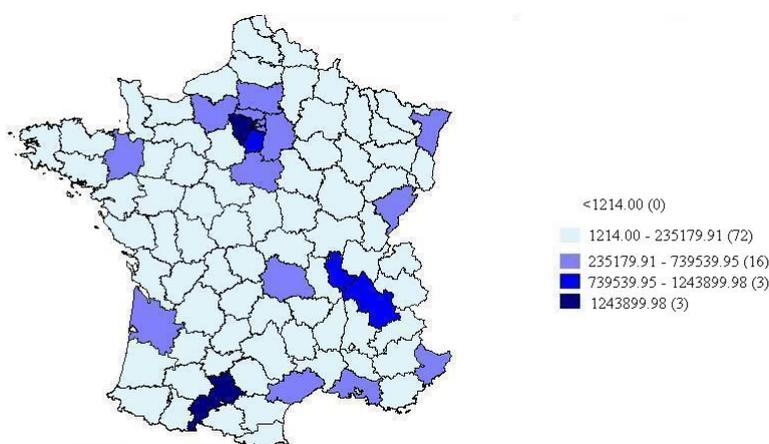


Figure 2.4 – Répartition des dépenses moyennes en R&D interne mesurées entre 2002 et 2005

- *La R&D externe privée et la R&D externe publique (derd)*

Contrairement à la R&D interne, le calcul de la R&D externe présente quelques difficultés dans la mesure où nous ne disposons pas de données par établissement. Pour ceci, nous avons construit un coefficient qui mesure la part de l'effectif R&D dans un établissement donné dans le total du même effectif calculé au niveau entreprise. Ce coefficient est multiplié par la suite par la R&D externe de chaque entreprise. Nous obtenons donc une valeur approximative de la R&D externe produite dans chaque établissement.

Ensuite, il nous paraît nécessaire de distinguer la R&D externe publique de la R&D externe privée. Aucun des travaux de la littérature ne semble faire cette distinction. La R&D externe publique regroupe la R&D confiée : à l'enseignement supérieur (universités, grandes écoles), au commissariat à l'énergie atomique (CEA), au centre français de recherche aérospatiale (ONERA), au centre national d'études spatiale (CNES), au centre national de la recherche scientifique (CNRS), aux institutions ou associations sans but lucratif faisant partie du secteur de l'état (hors centre technique (ISBL)).

La R&D externe privée est la somme des dépenses confiées aux 9 acteurs suivants : entreprises implantées sur le territoire français et appartenant au même groupe et au même secteur d'activité (ACT-GR), entreprises implantées sur le territoire français et n'appartenant pas au même groupe et au même secteur d'activité (AUT-ACT), entreprises implantées sur le territoire français et n'appartenant pas au même groupe et pas au même secteur d'activité (AUT-ENT), entreprises implantées sur le territoire français et appartenant au même groupe et pas au même secteur d'activité (AUT-GR), entreprises implantées à l'étranger et n'appartenant pas au même groupe (AUT-ETR), entreprises appartenant au même groupe mais à l'étranger (ETR-GR), organismes

2.5. Application de l'approche par cercles concentriques sur données en coupe

étrangers (ORG-ETR), organismes internationaux (ORG-INT), centres techniques professionnels (ORG-PRO) :

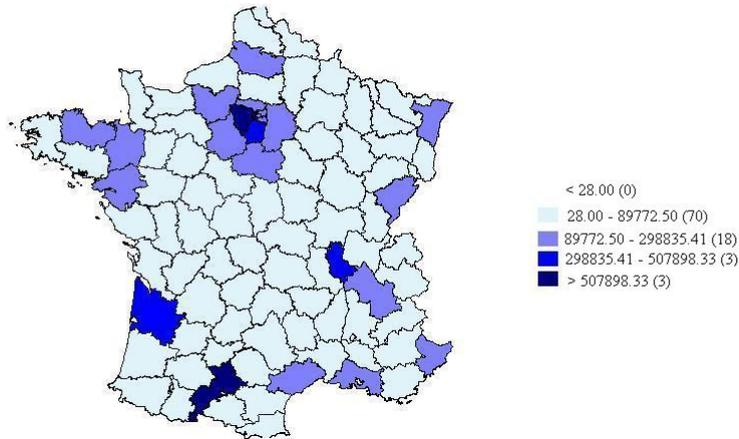


Figure 2.5 – Répartition des dépenses moyennes en R&D externe privée mesurées entre 2002 et 2005

Le montant d'investissement en R&D externe privée est de 87.657 milliers d'euros, soit 14 fois plus important que la R&D externe publique :

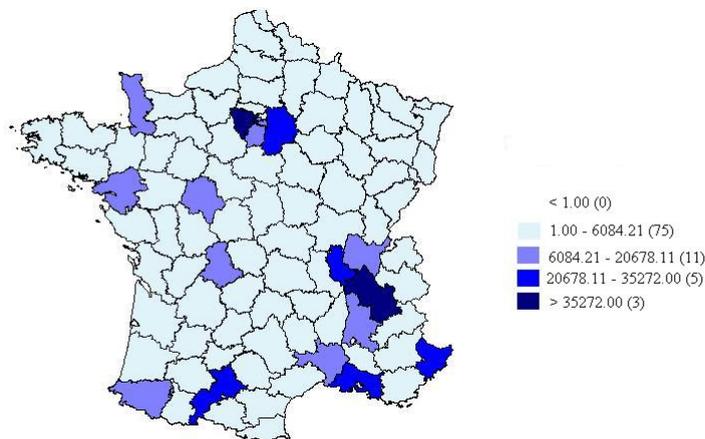


Figure 2.6 – Répartition des dépenses moyennes en R&D externe publique mesurées entre 2002 et 2005

Pôles : Nous construisons une variable "*pôle de compétitivité*" pour mesurer la quantité et la qualité des canaux de transmission des connaissances. Le pôle représente une combinaison d'entreprises et d'acteurs publics de la recherche et de la formation. Le comité interministériel d'aménagement et de compétitivité des territoires (CIACT) du 5 juillet 2007 a libellé 71 pôles de compétitivité. Parmi ces pôles, certains d'entre eux sont considérés comme des pôles mondiaux ou à vocation mondiale. Ces pôles sont novateurs en raison de leurs nombreux réseaux de collaboration de recherche et du type

des adhérents qui les composent, notamment la main d'oeuvre hautement qualifiée et mobile.

Dans le cadre des projets collaboratifs labellisés par les pôles, les entreprises et les organisations composant ces pôles bénéficient pour leurs projets collaboratifs de fonds supplémentaires financés par le FUI (Fond Unique Interministériel) et par l'ANR (Agence Nationale de la Recherche). Comme ils bénéficient d'un plus grand nombre de ces projets collaboratifs, nous supposons que les pôles mondiaux et à vocation mondiale sont plus efficaces dans le transfert des connaissances. Nous considérons donc que le nombre de pôle mondiaux ou à vocation mondiale présent dans un département est un bon indicateur synthétique de la qualité et de la quantité des canaux de diffusion des connaissances qui le caractérisent.

Pour construire cette variable proxy, nous allons prendre en compte le fait qu'un pôle mondial ou à vocation mondiale peut rassembler des entreprises situées dans plusieurs départements. Un département français peut accueillir plusieurs pôles dans différents domaines technologiques. Par exemple, deux pôles de classe mondiale ont été accordés au département Ile-et-Vilaine : un pôle spécialisé dans les TIC et le multimédia appelé "*pôle Images et Réseaux*", et un autre dédié aux activités de haute technologie liées à l'environnement maritime appelé "*Pôle Mer-Bretagne*".

Un pôle peut être aussi attribué à plusieurs départements en même temps selon le critère "effectif salarié". Par exemple, le pôle "*images et réseaux*" est attribué aux trois départements ayant l'effectif salarié le plus élevé : Ile-et-Vilaine, Loire-Atlantique et Finistère. L'effectif salarié le plus élevé du pôle "*Mer-Bretagne*" se trouve principalement dans les départements : Finistère, Morbihan et Ile-et-Vilaine. Nous nous retrouvons donc avec 29 départements dotés d'au moins un pôle mondial et à vocation mondiale. La localisation de ces pôles selon l'effectif est effectuée par la direction générale de l'industrie du ministère de l'Economie et des Finances. Le nombre moyen de pôles par département est de l'ordre de 0,54 et le nombre maximal est de 8. Tous ces pôles ont été accordés en 2005 sauf un en 2006 et un autre en 2007.

Capital Humain : Nous pouvons considérer aussi que ce n'est pas la quantité de la R&D qui importe, mais son intensité. Pour cela, nous construisons une variable mesurant le capital humain. Cette variable (noté KH) utilisée dans cette première partie empirique, rapporte le nombre de chercheurs de chaque département au nombre d'effectifs consacrés à la recherche dans ce même département. Elle mesure donc, pour chaque zone, la proportion de chercheurs par rapport à l'ensemble du personnel employé dans la recherche (chercheurs, techniciens, personnel administratif, etc.). L'effectif

chercheur de chaque département ainsi que le nombre de chercheurs sont issus de l'enquête R&D.

Taille de l'industrie : Cette variable exprime la taille moyenne en fonction de l'effectif de la R&D des établissements du secteur industriel d'un département i , rapporté à la taille moyenne des établissements de ce même secteur au niveau national.

$$Taille_{SecInd} = \frac{\frac{eff_i}{nb_i}}{\frac{eff_{ind}}{nb_{ind}}} \quad (2.8)$$

Avec eff désigne l'effectif de la R&D et nb le nombre d'établissements. i désigne le département et ind le secteur industriel.

Une influence négative de cette variable confirmerait le rôle important des firmes industrielles de petite taille dans les dynamiques d'externalités locales tandis qu'un rôle positif reviendrait à affirmer l'existence de rendements croissants de la recherche et devrait donc plutôt concerner les secteurs à fort coûts fixes de recherche (Massard et Riou, 2002).

Indice d'Ellison Glaeser : Afin de tenir compte des externalités de type Jacobs, nous utilisons deux variables destinées à cette fin. La construction de ces variables a exigé tout d'abord le choix d'un niveau d'agrégation des données industrielles. Nous avons opté pour 5 secteurs : l'industrie, le commerce, la construction, l'agriculture et les services et transports. Ensuite, nous avons utilisé l'effectif chercheur au lieu des dépenses en R&D pour construire cet indice de diversité industrielle.

Nous n'utilisons pas un simple indice d'Herfindhal puisque ce dernier ne tient pas compte de l'hétérogénéité de la dispersion des établissements dans les départements. Nous préférons un indice d'Ellison Glaeser (Ellison et Glaeser, 1997) calculé comme suit :

$$EGindex_i = -\frac{G_i - H_i}{1 - H_i} \quad (2.9)$$

$$\text{Avec } H_i = \sum_e \left(\frac{R\&D_e}{R\&D_i} \right)^2 \text{ et } G_i = \frac{\sum_k (S_{ik} - S_k)^2}{1 - \sum_k S_k^2},$$

où :

- S_{ik} est la part du nombre d'employés en R&D du secteur k dans le département i considéré,

- S_k est la part des employés en R&D du secteur k dans le nombre national d'employés en R&D,

- $R\&D_e$ est le nombre d'employés en R&D de l'établissement et,
- $R\&D_i$ le nombre d'employés en R&D dans le département i .

Les départements avec un indice $EGindex_i$ élevé sont caractérisés par une grande diversité de leurs activités de R&D. En revanche, ceux avec un indice faible sont caractérisés par des activités en R&D plus concentrées dans certains secteurs spécifiques. Un signe positif de l'indice d'Ellison Glaeser dans l'économétrie indique la présence d'externalités de type Jacobs.

Toutefois, il pourrait exister une corrélation positive entre la diversité cognitive d'un département et son degré d'urbanisation. Nous introduisons donc une autre variable afin de contrôler l'effet de l'urbanisation.

Urbanisation du département : Nous introduisons une variable dummy : "*largecity*" égale à 1 lorsque le département contient une ville de plus de 200.000 habitants.

Près de 12% des 94 départements de notre échantillon sont dans ce cas là.

Dans la table (2.1), les brevets, ainsi que les variables indépendantes sont divisées par la surface :

Tableau 2.1 – Statistiques descriptives pour les 94 départements français

	Min	Moyenne	Max	Ecart-type
$Bvttotal_i$	1	186,3085	1452	275,65
$R\&Dinttotal_i$	1214	235179,9	3109943	504360
$R\&Dextpivtotal_i$	28	87657,18	1311995	209400,9
$R\&Dextpubtotal_i$	1	95996	93519,52	14593,89
$KHtotal_i$	0,05	0,489	0,921	0,121
$Tailletotal_{i,sec2}$	0,02	0,699	5,689	0,747
$EGindex_i$	-3,307	-0,060	0,92	0,433
$largecity_i$	0	0,117	1	0,323
Bvt_i	0,0001	0,259	12,66	1,44
$R\&Dint_i$	0,032	522,647	57794,18	3615,547
$R\&Dextpiv_i$	0,0014	178,1376	18796,45	1276,805
$R\&Dextpub_i$	0,00005	12,198	1646,327	99,108
KH	9,68e-06	0,0002	0,007	0,0008
$Taille_{i,sec2}$	3,60e-06	0,00045	0,0165	0,002

$Bvttotal_i$: le nombre de brevets départemental entre 2006 et 2008,

$R\&Dinttotal_i$: l'investissement en R&D interne en millions d'euros entre 2002 et 2005,

$R\&Dextpiv_i$: l'investissement en R&D externe privée en millions d'euros entre 2002 et 2005,

$R\&Dextpub_i$: l'investissement en R&D externe publique en millions d'euros entre 2002 et 2005,

Bvt_i : nombre total de brevets au km².

En utilisant les variables présentées dans le tableau (2.1), nous estimons notre modèle en coupe en prenant en compte un décalage dans le temps de 3 ans. Ainsi, nous expliquons la moyenne du nombre de brevets au km² de chaque département, mesuré entre 2005 et 2008, par la moyenne de la R&D au km² de ce même département, mesurée entre 2002 et 2005. Nous supposons donc une structure de retard entre le moment où la R&D a eu lieu et le moment où elle débouche sur une invention (Autant-Bernard, 2000; Bound et al., 1984).

2.5.1.3. Le modèle

La spécification retenue est linéaire en logarithme. Nous retenons deux modèles en coupe : le premier est sans externalités géographiques (2.10) et le deuxième inclut des variables mesurant l'effet des externalités géographiques sur l'innovation (2.11).

$$\ln(Bvt_i) = \alpha + \beta_1 \ln(R\&Dint_i) + \beta_2 \ln(R\&Dextpub_i) + \beta_3 \ln(R\&Dextpiv_i) + \gamma_1 EGindex_{it} + \lambda_2 largecity_i + \lambda_3 poleM_i + \epsilon_i, \quad (2.10)$$

où :

- Bvt_i mesure le nombre moyen de brevets entre 2005-2008 au km^2 du département i et,
- $R\&Dint_i$ mesure les dépenses moyennes internes en R&D entre 2002-2005 au km^2 du département i ,
- $R\&Dextpub_i$ mesure la R&D externe moyenne publique entre 2002-2005 confiée à des organismes publics au km^2 du département i ,
- $R\&Dextpriv_i$ mesure la R&D externe moyenne privée entre 2002-2005 confiée à des organismes privés au km^2 du département i ,
- $EGindex_{it}$ est une mesure de la diversité industrielle moyenne du département i et pour la période 2005-2008,
- $poleM_i$: mesure le nombre de pôles mondiaux ou à vocation mondiale pour chaque département et pour la période 2005-2008,
- ϵ : le terme d'erreur.

En introduisant les externalités géographiques²¹, le modèle prend la forme suivante :

$$\begin{aligned} \ln(Bvt_i) = & \alpha + \beta_1 \ln(R\&Dint_i) + \beta_2 \ln(R\&Dextpub_i) + \beta_3 \ln(R\&Dextpriv_i) \\ & + \gamma_1 EGindex_{it} + \lambda_2 largacity_i + \lambda_3 poleM_i + \sum_{k=1}^3 [\beta_4 \ln(R\&D(k)_{moyen,]0-200]) \\ & + \beta_5 \ln(R\&D(k)_{moyen,]200-400]) + \epsilon_i \end{aligned} \quad (2.11)$$

Avec $k = 1, 2, 3$ et (1 : $R\&Dint_i$; 2 : $R\&Dextpub_i$ et 3 : $R\&Dextpriv_i$)

$-\ln(R\&D(k)_{moyen,]0-200])$ le logarithme des dépenses en R&D (interne, externe privée ou externe publique) mesurées dans un intervalle de 0-200km,

$-\ln(R\&D(k)_{moyen,]200-400])$ le logarithme des dépenses en R&D (interne, externe privée ou externe publique) mesurées dans un intervalle de 200-400km.

2.5.2. Estimations et résultats

Dans un premier temps, nous estimons un modèle sans effet de *spillovers* extra-départementaux. Dans un deuxième temps, nous introduisons les effets des *spillovers* mesurés dans les intervalles]0 – 200] et]200 – 400] puis dans les intervalles]0 – 100] et]100 – 200].

21. Nous avons aussi testé les effets de *spillovers* mesurés dans toutes les classes de (]0 – 100] jusqu'à l'intervalle [*plusde600kilomètre*[). Nous avons trouvé que les coefficients des variables mesurant les **spillovers** dans les intervalles au delà des 200 km sont tous non significatifs. Ceci rejoint le travail de Bottazi et Peri (2003) qui ont montré dans leur étude en coupe transversale sur la période 1977-1995, que les externalités de R&D existent seulement sur une distance de 300 kilomètres pour le cas des régions européennes (NUTS 2). Comme notre travail concerne un niveau plus fin (NUTS 3), l'étendue des *spillovers* est moins important dans ce cas.

Sans effet de *spillovers* extra-départementaux : Les résultats du modèle en coupe, sans effet de *spillovers* extra-départementaux (équation 2.10), sont reportés dans le tableau (2.2).

Nous expliquons le logarithme du nombre moyen de brevets au *km2*, calculé entre 2005 et 2008, par le logarithme des quantités de la R&D moyenne (R&D interne, R&D externe privée ou R&D externe publique) au *km2*, par la valeur moyenne de l'indice d'Ellison-Glaeser, par le degré d'urbanisation et par la variable pôle, calculés entre 2005 et 2008.

Tableau 2.2 – Résultats d'estimation des modèles sans externalités géographiques

$\ln(Bvt_{it})$	Coef	s.e	p-value
$\ln(R\&Dint_{t-1})$	0,666	0,038	0,000
$\ln(R\&Dextpriv_{t-1})$	0,009	0,003	0,013
$\ln(R\&Dextpub_{t-1})$	0,025	0,008	0,004
$EGindex_{it}$	0,229	0,125	0,069
$largacity_i$	-0,170	0,172	0,325
$poleM_i$	0,244	0,084	0,005
Constante	-6,426	0,114	0,000
R^2	0,913		
Test de Fisher	153,54		0,000
Test de Ramsey Reset (1)	6,79		0,0004
Test de Breusch-Pagan (2)	3,36		0,0666

s.e : Ecart-type, p-value : seuil de significativité.

(1) teste H_0 selon laquelle le modèle n'a pas de variables omises.

(2) teste H_0 selon laquelle la variance est constante.

Le test de Fisher est significatif au seuil de 1%, indiquant ainsi la significativité globale des variables explicatives. Par contre le test de variance des erreurs (test de Breusch-Pagan)²² indique un problème d'hétéroscédasticité (on rejette H_0 au seuil de 10%). En plus, le test de Ramsey (1969) est significatif indiquant le rejet de l'hypothèse nulle selon laquelle le modèle n'a pas de variables omises. Ces deux derniers tests révèlent une mauvaise spécification du modèle. C'est pourquoi, nous introduisons les effets de *spillovers* et nous vérifions la spécification du modèle avec *spillovers*.

Avec effet de *spillovers* extra-départementaux : Le tableau (2.3) présente les résultats des estimations avec effets de *spillovers*. Nous introduisons les mêmes variables indépendantes que celles introduites dans le modèle estimé ci-dessus, tout en rajoutant les effets de *spillovers* estimés dans les classes géographiques]0 – 200] et

22. Les régressions représentées dans les tableaux : (2.2), (2.3) et les 3 premières colonnes du tableau (2.4) sont établies sur stata en rajoutant l'option robust, afin de corriger les écarts types et minimiser le problème d'hétéroscédasticité.

]200–400]. Nous supposons ainsi que les connaissances et le savoir faire, représentés par les dépenses moyennes en R&D interne, R&D externe privée et R&D externe publique des départements j , peuvent être utilisées par les chercheurs du département i ($i \neq j$) :

Tableau 2.3 – Résultats d’estimation des modèles avec externalités géographiques mesurées dans des intervalles de 200 kilomètres

$\ln(Bvt_{it})$	Coef	s.e	p-value
$\ln(R\&Dint_{t-1})$	0,605	0,038	0,000
$\ln(R\&Dextpriv_{t-1})$	0,010	0,010	0,059
$\ln(R\&Dextpub_{t-1})$	0,029	0,017	0,096
$EGindex_{it}$	-0,159	0,123	0,201
$largacity_i$	0,011	0,192	0,06
$poleM_i$	0,242	0,061	0,000
$\ln(R\&Dint_{t-1,]0-200])$	-0,376	0,048	0,160
$\ln(R\&Dint_{t-1,]200-400])$	0,263	0,263	0,156
$\ln(R\&Dextpriv_{t-1,]0-200])$	0,004	0,004	0,007
$\ln(R\&Dextpriv_{t-1,]200-400])$	-0,002	0,223	0,998
$\ln(R\&Dextpub_{t-1,]0-200])$	0,120	0,078	0,128
$\ln(R\&DRDextpub_{t-1,]200-400])$	-0,002	0,233	0,993
Constante	-6.705	0.250	0,000
R^2	0,92		
Test de Fisher	82.12		0,000
Test de Ramsey Reset (1)	1,85		0,112
Test de Breusch-Pagan (2)	2,45		0,0028

s.e : Ecart-type, p-value : seuil de significativité.

(1) teste H_0 selon laquelle le modèle n’a pas de variables omises.

(2) teste H_0 selon laquelle la variance est constante.

Le test de Ramsey devient non significatif, indiquant l’acceptation de l’hypothèse nulle selon laquelle le modèle n’a pas de variables omises. Néanmoins, le test de Breusch-Pagan indique un problème d’hétéroscédasticité. En effet, la variance des erreurs n’est pas encore constante.

Les résultats des estimations (tableau 2.4) sont meilleurs lorsque nous estimons les *spillovers* dans des intervalles plus petits, en introduisant la variable taille de l’industrie (*TailleIndus*) et la variable *capital humain* (*KH*). Ces deux variables permettent d’instrumenter la variable pôle supposée endogène dans notre spécification (tableau 2.4, les trois dernières colonnes).

Nous établissons par la suite des tests (encadré 2.1) qui nous permettent de choisir entre les résultats des trois premières colonnes du tableau (2.4) (la variable *pôle* est

exogène) et les résultats des trois dernières colonnes du même tableau (la variable *pôle* est endogène).

Encadré 2.1 : Endogénéité et variables instrumentales

Considérons des modèles pour lesquels nous avons une équation linéaire entre une variable d'intérêt et des variables explicatives. Une variable explicative X est dite endogène si elle est corrélée avec le terme d'erreur : $Cov(x_t, \epsilon_t)$.

Nous distinguons trois sources principales d'endogénéité : (1) Erreurs de mesure; (2) Simultanéité et (3) Omission de variables. Cas (1) : si l'erreur de mesure sur la variable à expliquer est liée systématiquement à une ou plusieurs variables explicatives, cela peut entraîner un biais sur les estimateurs MCO.

Cas (2) : la simultanéité est la situation dans laquelle certaines des variables explicatives et la variable à expliquer sont déterminées "en même temps". Dans ce cas, l'estimateur MCO sera biaisé.

Cas (3) : l'omission de variables explicatives constitue une hétérogénéité inobservée. L'omission de cette variable conduit à une estimation non convergente du modèle par les MCO dès lors qu'elle est corrélée avec les explicatives.

L'endogénéité des variables explicatives peut être testée sur Stata avec le test d'Hausman.

Il existe aussi le test de Nakamura Nakamura qui se fait en deux étapes comme suit :

- Chaque variable endogène est régressé sur les variables exogènes du modèle et ses instruments.

- Les résidus de la première étape sont récupérés et inclus dans le modèle initial. Si les coefficients des résidus sont conjointement significatifs (test de Fischer) alors nous ne pouvons pas rejeter l'endogénéité des variables testées. Dans le cas d'une seule variable endogène, la significativité du t de student du résidu permet de conclure au non au rejet de l'hypothèse d'exogénéité. Une fois l'endogénéité est testée, une méthode des variables instrumentales est proposée. Cette méthode nécessite la connaissance de variables auxiliaires non corrélées avec les erreurs. Elle se ramène ensuite à l'application des MCO. Les instruments sont validés avec le test de suridentification de Sargan qui permet de tester la validité des instruments utilisés dans les régressions en double moindre carrés.

Les résultats trouvés après instrumentation de la variable *pôle* sont les suivants :

Tableau 2.4 – Résultats d’estimation des modèles avec externalités géographiques mesurées dans des intervalles de 100 kilomètres

$\ln(Bvt_{it})$	Coef	s.e	p-value	Coef	s.e	p-value
$\ln(R\&Dint_{t-1})$	0,626	0,049	0,000	0,531	0,123	0,000
$\ln(R\&Dextpriv_{t-1})$	0,011	0,010	0,003	0,007	0,010	0,011
$\ln(R\&Dpub_{t-1})$	0,024	0,017	0,007	0,018	0,025	0,548
$EGindex_{it}$	-0,07	0,135	0,444	-0,069	0,176	0,696
$largacity_i$	-0,128	0,195	0,511	-0,254	0,323	0,440
$poleM_i$	0,250	0,062	0,010	0,582	0,418	0,168
KH	0,938	0,719	0,214			
$TailleIndus$	-0,09	0,108	0,443			
$\ln(R\&Dint_{t-1,]0-100])$	0,088	0,036	0,018	0,103	0,049	0,042
$\ln(R\&Dint_{t-1,]100-200])$	0,090	0,477	0,849	-0,240	0,683	0,726
$\ln(R\&Dextpriv_{t-1,]0-100])$	0,007	0,013	0,008	0,008	0,014	0,009
$\ln(R\&Dextpriv_{t-1,]100-200])$	0,080	0,400	0,842	0,375	0,586	0,524
$\ln(R\&Dextpub_{t-1,]0-100])$	-0,002	0,027	0,931	0,002	0,035	0,951
$\ln(R\&Dextpub_{t-1,]100-200])$	-0,139	0,115	0,231	-0,100	0,139	0,476
Constante	-8,701	1,73	0,000	-7,803	1,965	0,000
R^2	0,9243			0,896		
(1)Test de Ramsey Reset	2,25		0,122			
(2)Test de Breusch-Pagan	2,51		0,113			
(3)Test de Dubin-Wu-Hausman				0,849		0,329
(4)Test de Sargan (Chi-sq)				1,284 (3)		0,2571(3)

s.e : Ecart-type.p-value : seuil de significativité

(1) teste H_0 selon laquelle le modèle n’a pas de variables omises.

(2)teste H_0 selon laquelle la variance est constante.

(3)basé sur la statistique de Fisher et teste H_0 selon laquelle la variable pôle est exogène.

(4) test de suridentification de Sargan-Hansen permet de tester H_0 selon laquelle les instruments utilisés ne sont pas corrélés avec les résidus.

La valeur critique du test de suridentification de Sargan²³ indique une bonne spécification du modèle. En d’autres termes, les variables instrumentales choisies ne sont pas corrélées avec les perturbations.

Les résultats du test de Dubin-Wu-Hausman après instrumentation nous conduit à accepter l’hypothèse nulle (l’exogénéité de la variable pôle). Nous concluons que la variable pôle n’est pas endogène. Nous retenons donc la première spécification du tableau (2.4) pour interpréter nos résultats.

23. Ce test n’est pas programmé sur Stata, il faut télécharger le module correspondant sur Internet. La commande utilisé est **overid** et fonctionne sous le même principe que la commande **ovtest**, mais la régression qui la précède ne doit pas contenir d’option **robust**.

Les estimations révèlent une forte élasticité des brevets par rapport à l'investissement en R&D interne : 0.626. L'élasticité des brevets par rapport aux dépenses en R&D externe privée (respectivement publique) est de l'ordre de 0.011 (respectivement 0.024). Les coefficients de la R&D externe sont faibles mais positifs et significatifs indiquant un effet positif de tout type d'investissement en R&D (interne ou externe) sur l'innovation. Ces résultats sont conformes à ceux trouvés sur des échantillons de données américaines et européennes (Jaffe et al., 1993 ; Feldman, 1994 ; Lung, 1997 ; Moreno et al., 2005).

Par ailleurs, l'effet R&D interne sur l'innovation est fortement positif et significatif. En effet, la R&D interne intègre plus de connaissances tacites, ce qui explique que son effet est beaucoup plus fort que celui de la R&D externe (contenant beaucoup plus des connaissances codifiées que tacites). La proposition 1 est donc validée.

L'indice d'Ellison Glaeser devient non significatif et nous ne pouvons pas conclure à travers cette dernière modélisation quant à la structure diversifié versus spécialisé de recherche dans les industries. Le coefficient de la variable urbanisation est positif mais non significatif. La variable pôle présente un coefficient positif et significatif indiquant un effet positif des canaux de transmission des connaissances sur la production d'innovation. En effet, les pôles mondiaux et à vocation mondiale présentent des réseaux riches en connaissances. Certains départements ont plus de clusters que les autres ce qui crée des différences en termes de capacité innovante. Nous validons ainsi notre proposition 2. Nous remarquons aussi que les résultats des estimations des effets de *spillovers* sont meilleurs lorsque nous mesurons ces effets dans des intervalles de 100km (en comparaison avec le tableau (2.3)).

Pour cela, nous nous limitons dans ce travail à montrer l'effet de ces externalités entre 0 et 200 kilomètres puisque au delà de cette distance, l'effet des externalités est presque inexistant. Lorsque nous introduisons les deux classes de distances pour chaque type de R&D, les résultats montrent que seulement les effets de *spillovers* géographiques de la R&D externe privée mesurée entre 0 et 100 kilomètres sont significatifs et positifs. L'élasticité des brevets à la R&D externe privée est de l'ordre de 0.007. Ainsi, si les dépenses en R&D externe privée mesurées dans]0 – 100km] augmenterait de 1%, les brevets dans le département considéré augmenterait de 0.007%.

Néanmoins, les données purement transversales utilisées dans cette section ne permettent pas de prendre en considération l'hétérogénéité entre départements. Autrement dit, certaines caractéristiques locales non prises en compte dans le modèle peuvent agir sur la capacité des départements à innover. Ces caractéristiques peuvent provenir par exemple des modalités d'organisation de la production ou de particularités

géographiques et culturelles de la zone (Autant-Bernard, 2000). Ceci constitue donc un risque d'obtention de résultats biaisés.

La prise en compte de l'hétérogénéité entre départements et la mise en oeuvre de méthodes d'estimation correspondantes devrait donc améliorer les résultats trouvés précédemment. Les données de panel présentent notamment l'intérêt de pouvoir contrôler l'hétérogénéité individuelle. En effet, un terme d'hétérogénéité régionale est pris en compte dans le modèle. Il capte les disparités non observées entre les départements qui peuvent affecter leurs capacités inventives²⁴.

2.6. Application de l'approche concentrique sur données de Panel

Nous essayons dans cette section d'appliquer la méthode de cercle concentrique sur des données de panel. Ces données présentent deux dimensions : une pour les individus i et une pour le temps t . Avec ces données, la question des effets individuels, de la corrélation et de l'hétéroscédasticité est ainsi traitée.

Nous analysons donc, la relation entre brevets et R&D en considérant les caractéristiques des 94 départements français entre 2002 et 2008. Cette double dimension, temporelle et géographique, permet d'utiliser des méthodes d'estimation plus efficaces, adaptées aux données de panel (Dormont, 1989).

2.6.1. Modèle et variables

Le modèle s'écrit alors comme suit :

$$\ln(Bvt_{it}) = \alpha + \beta_1 \ln(R\&Dint_{it-1}) + \beta_2 \ln(R\&Dextpub_{it-1}) + \beta_3 \ln(R\&Dextpiv_{it-1}) \\ + \gamma_1 EGindex_{it} + \gamma_2 EG0200_{it} + \lambda_1 networkscore_i + \lambda_2 largacity_i + \sum_t time_t + \\ \mu_i + \epsilon_{it}$$

où :

$-Bvt_{it}$ mesure le nombre total de brevets au km^2 du département i et pour l'année t :

24. Voir Baltagi (1995) pour une présentation plus générale des atouts et des inconvénients associés aux données de panel.

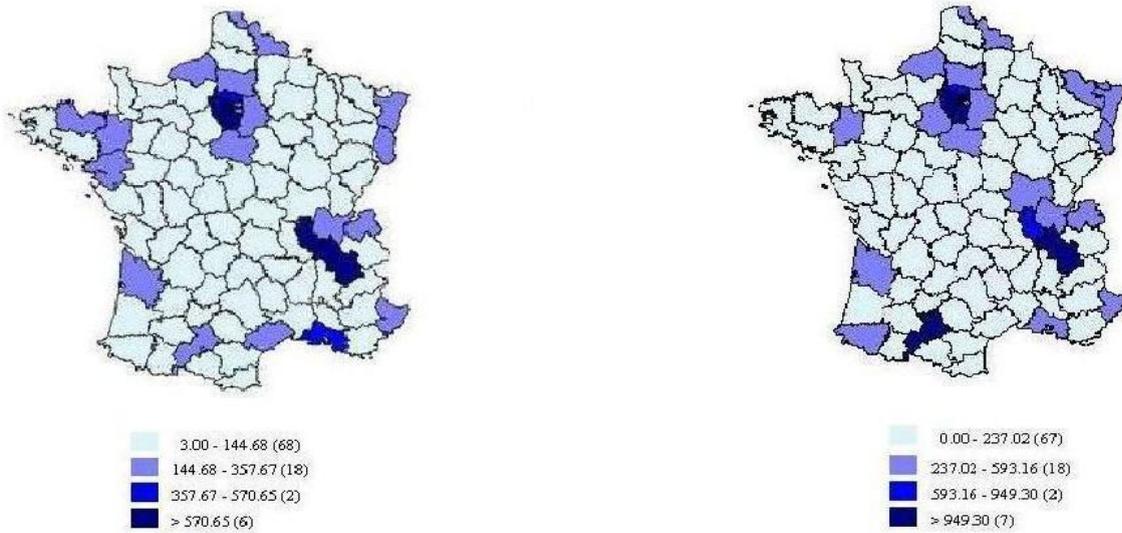


Figure 2.7 – Évolution du nombre des brevets entre 2002 et 2008

La distribution des brevets dans les départements n'a pas beaucoup changé entre 2002 et 2008. La figure (2.7) montre que les départements qui brevettent le plus sont toujours les mêmes (la région Ile de France, l'Isère ou le Rhône). Nous notons cependant une augmentation du nombre des brevets au sud de la France, notamment dans la Haute-Garonne.

- $R\&D_{int,t-1}$ mesure les dépenses internes en R&D au km^2 du département i et pour l'année $t-1$:

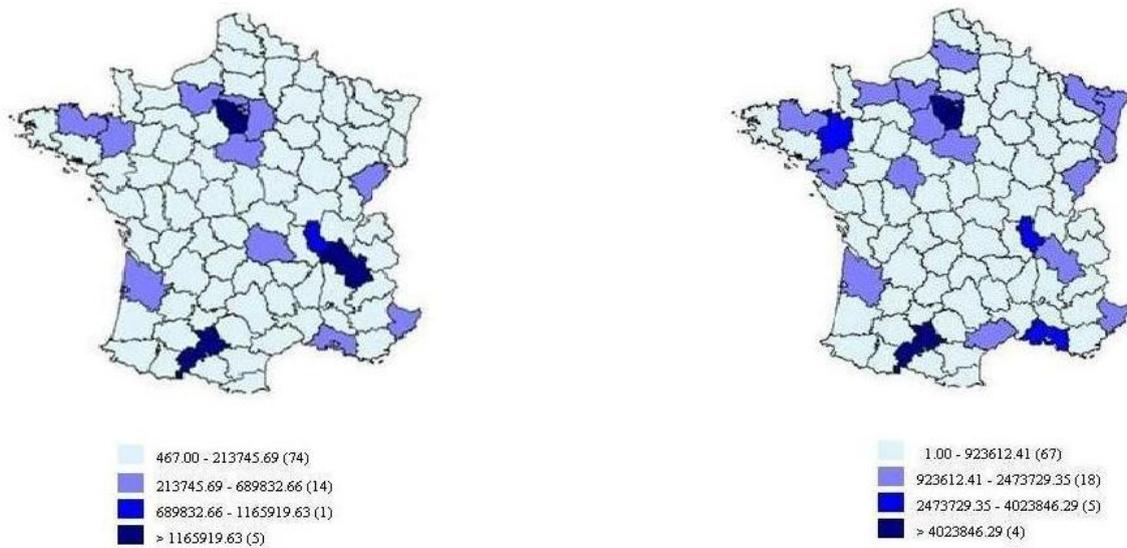


Figure 2.8 – Évolution des dépenses internes en R&D entre 2002 et 2008

$-R\&D_{extpub_{it-1}}$ mesure la R&D externe confiée à des organismes publics au km^2 du département i et pour l'année $t-1$:

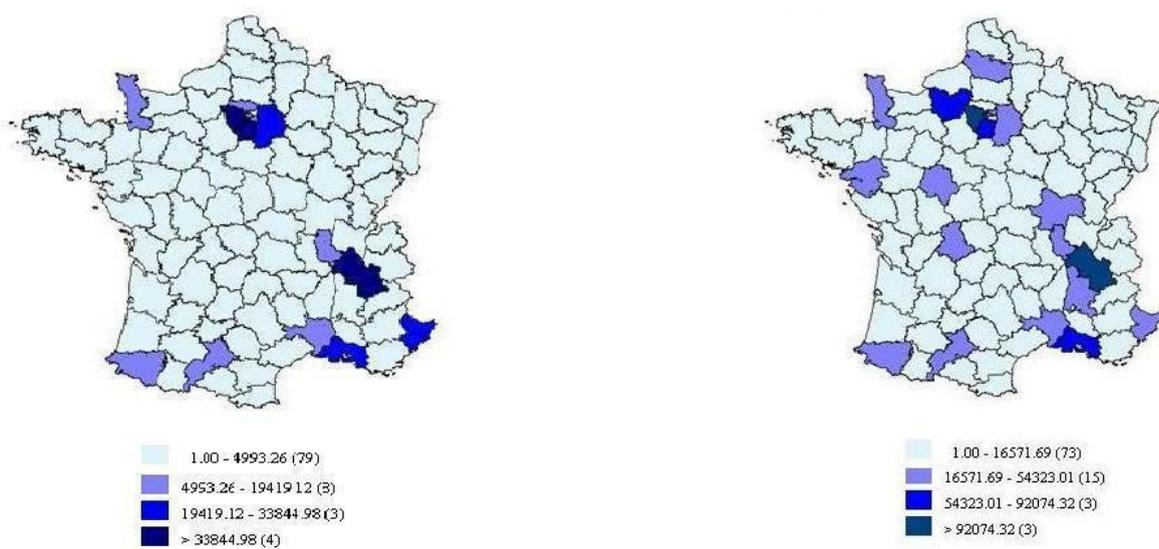


Figure 2.9 – Évolution des dépenses externes publiques en R&D entre 2002 et 2008

- $R\&Dextpriv_{it-1}$ mesure la R&D externe confiée à des organismes privés au km^2 du département i et pour l'année $t-1$:

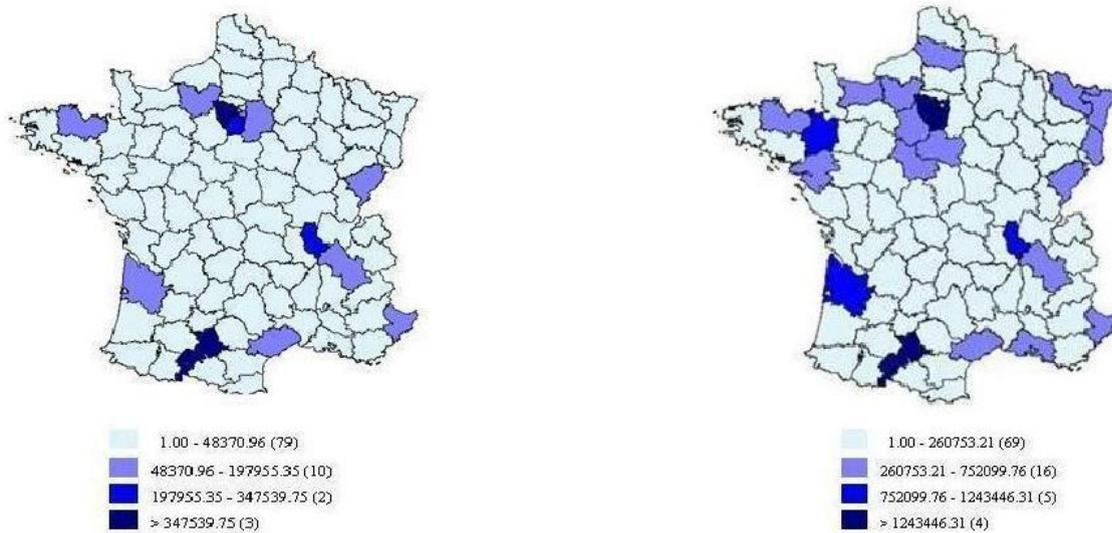


Figure 2.10 – Évolution des dépenses externes privées en R&D entre 2002 et 2008

La distribution de la R&D est aussi hétérogène que la distribution des brevets. La région Île-de-France occupe la première place en terme d'investissement en R&D (interne, externe privée et externe publique). Nous remarquons aussi une nette augmentation de ces dépenses pour la plupart des départements innovants entre 2002 et 2008.

- $EGindex_{it}$: est l'indice Ellison-Glaeser de diversité industrielle du département i et pour l'année t ,

- $EG0200_{it}$: est l'indice Ellison-Glaeser de diversité industrielle des départements voisins d'un département i , localisés dans un cercle de 0-200 kilomètres. Nous calculons le même indice pour les départements localisés dans un cercle de]0-100] kilomètres.

- $networkscore_i$: désigne l'une des variables suivantes $poleM_i, intercopubliscore_i, densityfuiscore_i$ ou bien $allnetworks_i$ définies comme suit :

- $poleM_i$: mesure le nombre de pôles mondiaux ou à vocation mondiale pour chaque département. Par exemple, si un département i possède 2 pôles en 2005, nous attribuons 2 au département considéré sur toute la période 2002-2008.
- $intercopubliscore_i$: est un score allant de 0 à 8. Cette variable est construite sur la base des déciles de la part moyenne des co-publications internationales dans les co-publications du département i entre 2002 et 2009.

- $densityfuiscore_i$: est un score allant de 0 à 8. Cette variable est construite sur la base des déciles mesurant la densité moyenne des réseaux FUI des pôles du département i .
- $allnetworks_i$: est la moyenne des variables $intercopubliscore_i$ et $densityfuiscore_i$.
- $time_i$: est une variable dummy égale à 1 à l'année $t=2004, \dots, 2008$; la variable temporelle représentant l'année 2002 est écartée (afin de tenir compte des problèmes de multicolinéarité entre variables), l'année 2003 est supprimée aussi puisque nous avons établi un décalage d'un an entre les dépenses en R&D et les brevets.
- μ_i : représente l'effet individuel non observable ;
- ϵ_{it} : le terme d'erreur.

L'analyse descriptive de ces variables est reportée dans le tableau (2.5) :

Tableau 2.5 – Statistiques descriptives des variables utilisées

Variable	Définition		Moyenne	Écart-type	Min	Max	Observations
$B_{ot,it}$	Nombre de brevets total au $km2$ du département i à l'instant t	Overall	175,7565	272,382	0,5	1953,256	N=658
		Between		258,037	1,833	1328,896	n=94
		Within		90,647	-210,949	993,175	T=7
$R\&D_{int,it}$	Dépenses en R&D interne au $km2$ du département i à l'instant $t-1$	Overall	432,755	2990,634	0,031	56192,03	N=564
		Between		2568,578	0,187	24111,25	n=94
		Within		150,77	-6868,6	32513,54	T=6
$R\&D_{extpub,it}$	Dépenses en R&D externe publique au $km2$ du département i à l'instant $t-1$	Overall	9,815	79,165	0,000	1278,26	N=564
		Between		69,553	0,002	670,166	n=94
		Within		38,371	-267,484	617,908	T=6
$R\&D_{privée,it}$	Dépenses en R&D externe privé au $km2$ du département i à l'instant $t-1$	Overall	154,611	1118,929	0,001	16801,64	N=564
		Between		955,486	0,036	9038,134	n=94
		Within		589,199	-4428,942	7918,119	T=6
$EG_{index,it}$	Indice d'Ellison-Glaeser : diversité sectorielle du département i à l'instant t	Overall	0,060	0,674	-4,113	11,416	N=658
		Between		0,433	-0,9195	3,307	n=94
		Within		0,518	-3,535	8,169	T=7
$EG0100_{,it}$	Le EG_{index} des départements voisins du département i à l'instant t et situés entre 0-100 km	Overall	0,328	5,017	-61,575	83,624	N=658
		Between		1,013	-5,981	3,900	n=94
		Within		4,914	-55,265	81,871	T=7

À Suivre

Variable	Définition	Moyenne	Écart-type	Min	Max	Observations
$EG0200_t$	Overall	-0,209	0,677	-4,478	0,600	N=658
	Between		0,542	-1,981	0,334	n=94
	Within		0,408	-0,707	1,799	T=7
$poleM_i$	Overall	0,543	1,156	0	8	N=658
	Between		1,161	0	8	n=94
	Within		0	0,543	0,543	T=7
$intercopubliscore_i$	Overall	1,65	2,710	0	8	N=658
	Between		2,723	0	8	n=94
	Within		0	1,65	1,65	T=7
$densityfuiscore_i$	Overall	1,681	2,732	0	8	N=658
	Between		2,744	0	8	n=94
	Within		0	1,681	1,681	T=7
$allnetworks_i$	Overall	1,665	2,556	0	7,5	N=658
	Between		2,568	0	7,5	n=94
	Within		0	1,665	1,665	T=7
$largecity_i$	Overall	0,117	0,321	0	1	N=658
	Between		0,323	0	1	n=94
	Within		0	0,117	0,117	T=7

La spécification retenue est celle du modèle à effets aléatoires. En effet, nous ne pouvons pas introduire les variables $PoleM_i$, $intercopubliscore_i$, $densityfuiscore_i$, $allnetworks_i$ et $largecity_i$ qui sont invariantes dans le temps dans un modèle à effet fixe. Par conséquent, les estimations à effets aléatoires sont les plus appropriées. Toutefois, les estimateurs à effets aléatoires ne régleraient pas le biais d'endogénéité potentiel généré par ces variables invariantes dans le temps si elles sont corrélées avec les effets individuels inobservés μ .

Pour résoudre ce problème, nous utilisons l'estimateur de Amemiya-MaCurdy (1986). Cet estimateur est une extension de celui d'Hausman-Taylor (1981) pour une spécification à effets aléatoires avec corrélation entre variables et effets individuels non observés. Il suppose, cependant, que ces variables explicatives sont indépendantes du terme d'erreur ϵ_{it} .

L'estimation de notre modèle avec Hausman-Taylor ne peut fournir des estimateurs convergents et efficaces que si les instruments utilisés respectent certaines conditions de validité. La méthode d'Amemiya MaCurdy est donc plus efficace dans la mesure où elle place des restrictions sur l'ensemble des instruments. Elle permet donc d'améliorer l'efficacité de l'estimateur (Sevestre, 2002).

Nous validerons l'hypothèse selon laquelle les variables explicatives instrumentées sont indépendantes du terme de l'erreur avec le test de suridentification de Sargan-Hansen.

2.6.2. Estimations et résultats

Le tableau suivant présente les résultats d'estimation avec Amemiya-MaCurdy. Les résultats affichés sont obtenus en utilisant quatre variables alternatives mesurant la qualité et la quantité des réseaux de recherche : $densityfuiscore_i$, $intercopubliscore_i$, $allnetworks_i$, $poleM_i$.

Tableau 2.6 – Résultats d’estimation de la FPI sur des données de panel sans *spillovers* extra-départementaux

	$\ln(Bvt_{it})$	$\ln(Bvt_{it})$	$\ln(Bvt_{it})$	$\ln(Bvt_{it})$
$\ln(R\&Dint_{it-1})$	0,131**** (0,030)	0,131**** (0,030)	0,133**** (0,030)	0,143**** (0,031)
$\ln(R\&Dextpub_{it-1})$	0,022* (0,011)	0,022* (0,011)	0,022* (0,011)	0,021* (0,012)
$\ln(R\&Dextpriv_{it-1})$	0,021 (0,014)	0,021 (0,014)	0,021 (0,014)	0,022 (0,014)
$EGindex_{it}$	0,108**** (0,025)	0,108**** (0,025)	0,108**** (0,025)	0,108**** (0,026)
$EG0200_{it}$	0,053* (0,028)	0,053* (0,028)	0,053* (0,028)	0,051* (0,029)
$densityfuiscore_i$	0,521** (0,181)			
$intercopubliscore_i$		0,443** (0,199)		
$allnetworks_i$			0,521*** (0,184)	
$poleM_i$				1,292**** (0,200)
$largecity_i$	-0,135 (1,034)	-0,272 (1,216)	-0,365 (1,058)	-0,397 (0,544)
<i>Constante</i>	-5,371**** (0,384)	-5,211**** (0,398)	-5,340**** (0,371)	-5,195**** (0,189)
Observations	564	564	564	564
Sargan	11,867 Chi-sq(29) p-value = 0,9979	12,930 Chi-sq(29) p-value = 0,9956	13,714 Chi-sq(29) p-value = 0,9927	20,531 Chi-sq(29) p-value = 0,8757
Wald Chi2 test ^a	10,40***	10,60***	10,75***	12,55****
Wald Chi2 test ^b	9,33***	9,52***	9,62***	10,79***

Les variables instrumentées sont respectivement : $densityfuiscore_i$, $intercopubliscore_i$, $allnetworks_i$, $poleM_i$.

^a La statistique de χ^2 testant l’hypothèse selon laquelle la différence entre les effets marginaux des variables $\ln(R\&Dint_{it-1})$ et $\ln(R\&Dextpub_{it-1})$ est nulle.

^b La statistique de χ^2 testant l’hypothèse selon laquelle la différence entre les effets marginaux des variables $\ln(R\&Dint_{it-1})$ et $\ln(R\&Dextpriv_{it-1})$ est nulle.

Les niveaux de significativité statistique sont indiqués par : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

Les résultats des estimations reportés dans le tableau (2.6), confirment l'effet positif des variables R&D sur l'innovation (résultat trouvé précédemment avec des données en coupe). En effet, les variables internes au département i importent dans le processus d'innovation. Ces résultats sont cohérents aussi avec ceux trouvés par d'autres études sur les départements français (Autant-Bernard, 2001a,b; Massard et Riou, 2002; Autant-Bernard et LeSage, 2011) et aussi avec les études faites au niveau des régions européennes (Bottazzi et Peri, 2003; Fischer et Varga, 2003; Ponds et al., 2010).

La statistique de Sargan-Hansen montre que les instruments sont valides dans toutes les régressions. L'élasticité de la R&D par rapport aux brevets se situe entre 0,13 et 0,14²⁵ et est fortement significative dans toutes les spécifications.

La R&D externe publique est légèrement significative avec une élasticité de 0,022, mais pas la R&D externe privée.

Nous constatons aussi que la R&D interne a un effet plus important que celui de la R&D externe publique et privée sur l'innovation. Les tests de Wald rejettent l'hypothèse selon laquelle la différence entre les coefficients est nulle.

L'effet de la R&D interne est six fois plus important sur l'innovation que celui de la R&D externe (publique et privée). En effet, la R&D interne, contrairement à la R&D externe, intègre davantage de connaissances tacites. Elle intègre aussi des connaissances codifiées, mais la R&D externe n'a pas elle de connaissances tacites, ou très peu. Ce résultat confirme notre proposition 1.

L'indice d'Ellison-Glaeser est fortement significatif dans toutes les estimations avec un signe positif indiquant qu'une structure de recherche plus diversifiée est bénéfique à la productivité de l'innovation départementale. L'indice d'Ellison-Glaeser des régions voisines dans le cercle 0-200 km est également positif et significatif mais avec un coefficient divisé par deux. Ainsi, la diversification dans la périphérie d'un département i a un effet positif sur l'innovation de ce même département, et est significativement positive et efficace même au delà des frontières départementales. Ce dernier résultat nous permet de valider notre proposition 3.

Les variables instrumentées : $densityfuiscore_i$, $intercopubliscore_i$, $allnetworks_i$, $poleM_i$ sont toutes fortement significatives avec les signes positifs attendus, ce qui confirme notre proposition 2.

25. Le coefficient de la R&D interne est passé de 0,54 dans le tableau (2.4) (estimation en coupe) à 0,13 et 0,14 (estimation en panel). Cette différence est expliquée par la prise en compte de l'hétérogénéité entre individus. Ainsi, l'observation répétée dans le temps sur les mêmes individus permet de mieux comprendre les différences entre individus et l'évolution temporelle des phénomènes étudiés.

La densité des réseaux constitués par les collaborations FUI régionales est positivement corrélée avec la productivité de l'innovation. Le coefficient de la variable $intercopubliscore_i$ mesurant la part des co-publications internationales dans co-publications des pôles mondiaux et à vocation mondiale des départements est aussi positif et significatif.

Le coefficient de la variable $allnetworks_i$ est moins important que celui de la variable $densityfuiscore_i$, mais par contre, celui de la variable $poleM_i$ (qui mesure la quantité et de la qualité de tous les réseaux de recherche régionaux), affiche un coefficient beaucoup plus élevé (1.292 au lieu de 0.521), ce qui semble valider l'idée selon laquelle l'effet des réseaux est cumulatif (proposition 2).

Cependant, ce résultat doit être considéré avec prudence puisque la variable $PoleM_i$ pourrait toujours capter d'autres caractéristiques départementales, autre que les réseaux, même si nous pensons que nous avons contrôlé la plupart de ces caractéristiques.

Au regard des résultats trouvés, l'analyse a permis de confirmer les hypothèses présentées dans la section 3 de ce chapitre. Néanmoins, nous ne pouvons pas déterminer les étendues géographiques des effets des externalités de la R&D (interne et externe) sur l'innovation sans les introduire dans la FPI. Pour cela, nous introduisons dans notre équation les externalités géographiques de R&D mesurées comme suit :

$$\begin{aligned}
 \ln(Bvt_{it}) = & \alpha + \beta_1 \ln(R\&Dint_{it-1}) + \beta_2 \ln(RDint_{t-1,]0-100]) + \beta_3 \ln(R\&Dint_{t-1,]100-200]) \\
 & + \beta_4 \ln(R\&Dextpub_{it-1}) + \beta_5 \ln(RDextpub_{t-1,]0-100]) + \beta_6 \ln(R\&Dextpub_{t-1,]100-200]) \\
 & + \beta_7 \ln(R\&Dextpriv_{it-1}) + \beta_8 \ln(R\&Dextpriv_{t-1,]0-100]) + \beta_9 \ln(R\&Dextpriv_{t-1,]100-200]) + \\
 & \gamma_1 EGindex_{it} + \gamma_2 EG0100_{it} + \lambda_1 networkscore_i + \lambda_2 largacity_i \\
 & + \sum_t time_t + \mu_i + \epsilon_{it} \tag{2.12}
 \end{aligned}$$

Les résultats de l'application de l'approche par cercles concentriques sur données de panel avec introduction des externalités de R&D sont reportés dans le tableau (2.7) :

Tableau 2.7 – Résultats d'estimation de la FPI sur des données de panel avec *spillovers* extra-départementaux.

	$\ln(Bvt_{it})$	$\ln(Bvt_{it})$
$\ln(R\&Dint_{it-1})$	0.140**** (0.030)	0.160**** (0.030)
$\ln(R\&Dextpub_{it-1})$	0.021* (0.011)	0.021* (0.011)

2.6. Application de l'approche concentrique sur données de Panel

$\ln(R\&Dextpriv_{it-1})$	0.020 (0.013)	0.023 (0.014*)
$EGindex_{it}$	0.100**** (0.024)	0.099**** (0.025)
$EG0100_{it}$	0.001 (0.0027)	0.002 (0.002)
$intercopubliscore_i$	0.0449**** (0.019)	
$poleMi$		0.880**** (0.139)
$largecity_i$	0.406 (0.87)	0.154 (0.468)
$\ln RDint_{t-1,]0-100]}$	0.11** (0.056)	0.118**** (0.053)
$\ln RDint_{t-1,]100-200]}$	-0.242**** (0.097)	-0.219** (0.094)
$\ln RDextpub_{t-1,]0-100]}$	0.034 (0.034)	0.033 (0.034)
$\ln RDextpub_{t-1,]100-200]}$	-0.167**** (0.050)	-0.170**** (0.050)
$\ln RDextpriv_{t-1,]0-100]}$	-0.039 (0.034)	-0.035 (0.035)
$\ln RDextpriv_{t-1,]100-200]}$	0.226**** (0.061)	0.234**** (0.062)
<i>constante</i>	-4.342**** (1.636)	-4.758**** (1.546)
Observations	564	564
Sargan	0.002 Chi-sq(1) p-value = 0.9652	0.055 Chi-sq(1) p-value = 0.8145
Wald Chi2 test ^a	15.60****	20.05***
Wald Chi2 test ^b	14.83****	17.99***

Les niveaux de significativité statistique sont indiqués par : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

^a La statistique de χ^2 testant l'hypothèse selon laquelle la différence entre les effets marginaux des variables $\ln(R\&Dint_{it-1})$ et $\ln(R\&Dextpub_{it-1})$ est nulle.

^b La statistique de χ^2 testant l'hypothèse selon laquelle la différence entre les effets marginaux des variables $\ln(R\&Dint_{it-1})$ et $\ln(R\&Dextpriv_{it-1})$ est nulle.

Les résultats du tableau (2.7) affichent des coefficients similaires à ceux trouvés dans le tableau (2.6) (pour $\ln(R\&Dint_{it-1})$, $\ln(R\&Dextpriv_{it-1})$, $\ln(R\&Dextpub_{it-1})$ et pour $EGindex_{it}$).

Cependant, les effets des *spillovers* sont différents de ceux trouvés avec une spécification en coupe (tableau 2.4) : nous trouvons des effets à la fois positifs et négatifs.

Le tableau (2.7) montre que les externalités de la R&D externe privée mesurées dans l'intervalle]0 – 100km] ont un effet non significatif et négatif sur l'innovation du département considéré. L'effet de ces externalités mesurées dans]100 – 200km] devient significatif et positif.

Cependant, les effets des externalités de la R&D interne et des externalités de la R&D externe publique sont différents : l'effet des externalités de la R&D interne mesurées dans l'intervalle]0 – 100km] est positif et significatif. Par contre, cet effet devient fortement négatif et significatif dans l'intervalle]100 – 200km]. Concernant les externalités de la R&D externe publique, leur effet est non significatif dans l'intervalle]0 – 100km], mais devient significatif et négatif dans]100 – 200km].

Abdelmoula et Bresson (2007), qui ont mesuré l'effet des externalités de R&D sur l'innovation au niveau des régions européennes, ont donné une explication à cette variation de signe dans les intervalles. Selon ces auteurs, les effets de *spillovers* négatifs (respectivement positifs) peuvent être associés à une concurrence (respectivement complémentarité) entre régions européennes.

Pour notre cas, les départements très proches (dans un rayon de 100km) sont complémentaires (pour la R&D interne et la R&D externe publique) tandis que les départements distants sont rivaux. Les départements se complètent dans la R&D interne et externe publique jusqu'à 100km et se concurrencent dans le segment]100-200].

Ces résultats vont dans le sens des apports théoriques de Boschma (2004), Boschma et Lambooy (1999) et Boschma et al. (2002) qui présentent la proximité géographique comme une contrainte. Ils soutiennent que trop de proximité géographique peut être néfaste pour l'innovation.

Au final, nous voyons que les résultats de modélisation d'une FPI avec données en coupe sont différents de ceux trouvés avec des données en panel. Même si l'application de la méthode des cercles concentriques sur des données de panel a permis de tenir compte de l'hétérogénéité entre les observations, nous ne pouvons pas conclure quant à l'efficacité inhérente de cette méthode.

En effet, cette méthode inspirée de Bottazzi et Peri (2003) se base sur les centroïdes des départements (latitude et longitude) qui nous permettent de déterminer la localisation géographique de chaque département. En revanche, si une grande partie de la surface d'un département se trouve à l'extérieur du rayon 100 km par exemple et que son centroïde est à l'intérieur, la totalité de la R&D sera attribué au rayon 0-100.

La mesure de toutes les variables a été réalisé de cette manière. Ceci représente une limite majeure de cette méthode des cercles concentriques basée sur les centroïdes des départements français. Malheureusement, nous ne disposons pas de données à des niveaux géographiques plus désagrégés comme par exemple les cantons qui permettraient de contourner ce problème rencontré avec des données départementales.

2.7. Conclusion

Deux principales approches empiriques sont mobilisées dans la littérature pour étudier la relation entre proximité et innovation par le biais des "*spillovers*". Une première mesure les flux de connaissance en utilisant les citations de brevets. Une deuxième s'appuie sur la Fonction de Production de Connaissances de type Jaffe (1989). Cette dernière est la plus utilisée dans la littérature empirique sur les mesures des *spillovers* puisqu'elle présente l'avantage de mesurer l'interaction entre unités voisines et proches (Autant-Bernard, 2001a,b; Bottazzi et Peri, 2003).

En nous inspirant des travaux de Bottazzi et Peri (2003), nous avons estimé l'effet des externalités de R&D sur le processus d'innovation. Nous avons décomposé la R&D en R&D interne, R&D externe publique et R&D externe privée et nous avons estimé l'effet de leurs *spillovers* sur l'innovation des 94 départements français métropolitains.

Dans un premier temps, nous avons utilisé comme Bottazzi et Peri (2003) des données en coupe et nous avons testé l'effet des *spillovers* entre 2002 et 2005 sur l'innovation des départements français, mesurée par le nombre moyen de brevets entre 2005 et 2008. Nous avons introduit dans le modèle plusieurs variables mesurant la qualité et la quantité des réseaux d'innovation dans les départements. L'objectif était double : déterminer l'étendue géographique de ces externalités et montrer l'importance des réseaux d'innovation dans la diffusion des connaissances et des externalités. Dans un deuxième temps, nous avons appliqué la même méthode de cercles concentriques sur des données de panel pour une période allant de 2002 à 2008. Nos résultats sont divergents en ce qui concerne l'effet et les étendues des externalités de R&D.

Dans l'étude en coupe, nous avons obtenu que seulement l'effet des externalités de la R&D externe privée mesurées dans 0-100 km est significatif et positif. Dans les autres intervalles, l'effet des externalités est presque inexistant. Ce résultat montre que

le transfert des connaissances codifiées ne nécessite pas trop de proximité géographique, contrairement aux connaissances tacites (mesurées par la R&D interne) qui nécessitent beaucoup plus de contacts face à face.

Les résultats trouvés avec une étude sur données de panel sont différents. En effet, les effets des *spillovers* sont à la fois négatifs et positifs selon le type de la R&D et l'intervalle considéré. L'approche des cercles concentriques présente donc certaines limites qui font que les résultats trouvés avec des données en coupe sont différents de ceux trouvés avec données de panel. C'est pourquoi nous associons dans le chapitre suivant ces méthodes à des techniques de l'économétrie spatiale afin de mesurer de manière plus fine l'ampleur des effets d'interdépendance dans l'espace.

En effet, lorsqu'il est difficile de modéliser à priori l'étendue spatiale des externalités (Autant-Bernard et al., 2007), l'économétrie spatiale constitue une solution en permettant de repérer dans les données, les phénomènes d'interactions entre zones ou entre agents situés, puis de prendre en compte ces phénomènes dans les estimations (Anselin, 2003).

Nous présenterons dans le chapitre 3 les outils de détection de l'autocorrélation spatiale locale. L'apport de la modélisation de la FPI par les outils de l'économétrie spatiale est présenté en détail dans le chapitre 4.

Chapitre 3

Diffusion spatiale des connaissances en France : Une analyse exploratoire de données spatiales

3.1. Introduction

En dépit de la diversité des méthodes utilisées pour mesurer l'étendue géographique des "*externalités de connaissance*", aucune d'entre elles ne semble exempte de critiques. Ces méthodes, notamment la fonction de production d'innovation (FPI), ont été discutées dans le chapitre 2 de notre thèse. D'ailleurs, nous avons montré que l'utilisation de la FPI avait deux objectifs principaux. Le premier consistait à déduire la présence des externalités de connaissance à travers la mesure de la concentration des activités innovantes (ou en introduisant des indices de coïncidences géographiques dans la FPI). Le deuxième s'intéressait surtout à la mesure de l'étendue spatiale de ces externalités (en mesurant les externalités de R&D dans des intervalles de distances géographiques différents). Néanmoins, nous avons vu que l'utilisation de la FPI pour modéliser les externalités a posé un certain nombre de problèmes liés à la nature complexe de ces dernières et à la difficulté de leur modélisation.¹

Les externalités de connaissance ont constitué, comme nous l'avons montré dans le chapitre premier, une explication plausible de la forte concentration des activités innovantes. Même si nous n'avons pas de preuve empirique définitive de l'existence de ces externalités, c'est-à-dire que nous ne pouvons pas dire s'il s'agit vraiment d'externalités de connaissance ou d'une simple diffusion des connaissances, tous les résultats empiriques montrent la concentration des activités innovantes dans l'espace.

1. Notamment les problèmes liés à la confusion entre diffusion des connaissances et externalités.

Toutefois, même si l'intensité de la concentration est considérée comme un symptôme de l'existence d'externalités au niveau local, elle n'en est pas en soi une preuve suffisante.

Ainsi, nous explorons dans ce chapitre les techniques de l'autocorrélation spatiale permettant d'analyser la structure de concentration des activités d'innovation (Guillain et le Gallo, 2007). Ces techniques permettent d'étudier les caractéristiques de la concentration d'un point de vue *intensité* et *structure* des activités innovantes. Au delà de la simple observation de la concentration géographique, une analyse par les techniques de l'autocorrélation spatiale permet de détecter l'existence de d'interactions entre deux unités spatiales par exemple, en apportant de nouveaux éclairages sur la nature de la causalité entre les observations. Ces interactions spatiales entre observations seront prises en compte par la suite dans les estimations empiriques du chapitre 4.

Notre chapitre est divisé en trois sections. Dans la section 2, nous analysons la concentration des brevets et des dépenses en R&D pour les années 2002 et 2008. La section 3 est consacrée à l'étude de l'intérêt d'une prise en compte de la dépendance spatiale à travers l'utilisation des outils de l'autocorrélation spatiale globale. Dans la section 4, nous exploitons les outils d'analyse de l'autocorrélation spatiale locale afin d'identifier les *clusters* innovants en France. Enfin, la section 5 conclut le chapitre.

3.2. Concentration et agglomération de l'innovation en France

La concentration géographique est définie comme la distribution du poids des départements dans un secteur spécifique de l'activité économique j . Dans ce sens, une activité spécifique j est considérée "*concentrée*" si une grande partie de la production est réalisée dans un nombre réduit de départements (Aiginger et al., 1999).

Nous analysons dans cette section le degré de concentration des activités innovantes en France et leurs schémas de localisation. Nous montrons par la suite, la nécessité de l'utilisation des outils de mesure de l'agglomération pour examiner la structure spatiale des localisations.

3.2.1. Mesure de la concentration des brevets et des dépenses en R&D en France

Le point de départ en géographie de l'innovation est l'étude de la concentration spatiale des activités de R&D ou d'innovation utilisée comme une preuve empirique des externalités de connaissance localisées.

Afin de mieux comprendre les effets de diffusion à la base des externalités locales, nous analysons dans cette section la concentration des activités innovantes au niveau des départements français. L'étude de la concentration de l'innovation a été initiée par Feldman (1994a) et Audretsch et Feldman (1996) dans le contexte Américain, Moreno, Paci et Usai (2005) en Europe et Autant-Bernard (2000), Gallié (2004) et Guillain et le Gallo (2007) en France. En calculant l'indice de Gini, ces travaux montrent que l'innovation est marquée par une forte polarisation spatiale, une concentration qui est même plus importante que les autres activités de production. Aux Etats-Unis, l'essentiel de ces innovations est concentré sur les côtes, dans les Etats de Californie, New York et New Jersey. En Europe, la moitié de l'effort de recherche est le fait de trois pays seulement : Allemagne, France et Royaume-Uni. Plus de 81% des brevets de l'OCDE sont déposés par des inventeurs localisés dans des régions urbaines. La place de ces zones urbaines est particulièrement marquée aux Pays-Bas (95%), en Belgique (88%) et aux Etats-Unis (78%). Autant-Bernard (2000) montre en utilisant des données multi-sectorielles et des indicateurs de l'innovation (brevets, R&D et publications) au niveau départemental, que la concentration des activités de recherche et d'innovation est supérieure à celle de la localisation des activités de production industrielles. Ce résultat est conforté par les indices de concentration des indicateurs de l'innovation pondérés par la valeur ajoutée de chaque département. Gallié (2004) aboutit à la même conclusion en étudiant la concentration des activités d'innovation dans le secteur des biotechnologies en France. Ensuite, Guillain et Le Gallo (2006) ont montré dans le cas de la région Île-de-France, que certains secteurs comme la R&D et celui de la finance et de l'assurance ont tendance à se concentrer au niveau de *clusters*.

A partir de ces études, nous remarquons que les brevets et les dépenses en R&D ont, en effet, tendance à se regrouper en certains points particuliers de l'espace, ce qui conduit à une concentration géographique des activités innovantes². Ce "*fait stylisé*" (Easterly et Levine, 2001) se retrouve à tous les niveaux de l'échelle spatiale : au niveau mondial, national ou régional. Ces études empiriques récentes sur la localisation des activités économiques cherchent donc, à travers l'utilisation des indicateurs adéquats, à mesurer la concentration géographique et à expliquer en même temps comment celle-ci se forme et s'accroît. L'indice le plus utilisé est celui de Gini, mais il en existe d'autres, notamment l'indice d'Herfindahl et l'indice d'Ellison et Glaeser. L'indice de Gini a d'abord été utilisé pour mesurer des inégalités de revenus entre individus. Il fut appliqué par la suite à l'économie spatiale en termes de mesure d'une activité économique donnée : emploi, production, innovation, etc., et d'un secteur donné.

2. Les cartes de l'évolution des nombre de brevets ainsi que celles de l'évolution des dépenses en R&D (interne, externe publique et externe privée) ont été présentées dans la sous section 2.6.1 du chapitre 2.

Par ailleurs, cet indicateur de la concentration des activités économiques ne donne pas d'indication sur la nature diversifiée ou spécialisée des activités concentrées³. Au niveau des départements français, la concentration de l'innovation est très remarquable pour les brevets, les dépenses en R&D interne, les dépenses en R&D externe publique et les dépenses en R&D externe privée en 2002 et 2008⁴.

En 2002, la courbe de Lorenz est représentée comme suit :

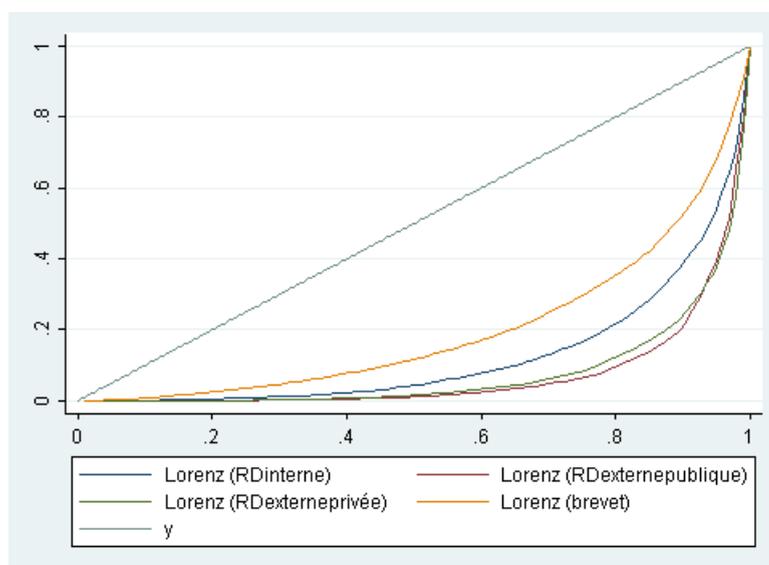


Figure 3.1 – Concentration des brevets et de la R&D interne et externe en 2002

En 2008, nous remarquons une légère baisse de la concentration des dépenses en R&D et une augmentation de la concentration des brevets :

3. Comme nous l'avons précisé dans le chapitre 2, l'indice d'Ellison et Glaeser permet d'examiner l'effet de la variété et de la diversité des secteurs (Externalités Jacobs) versus la spécialisation sectorielle locale (Externalités MAR) sur l'innovation. L'objectif de cette section est de mesurer la concentration globale. Nous nous limitons donc à l'utilisation de l'indice de Gini qui mesure la concentration de l'innovation tout en représentant graphiquement la disparité départementale à travers la courbe de Lorenz.

4. Nos données sont estimées pour la période 2002-2008. Pour cela, nous avons choisi de comparer la concentration de l'innovation et de son agglomération au niveau des départements français pour les deux dates suivantes : 2002 (représentant le début de la période d'étude considérée) et 2008 (représentant la fin de la période d'étude considérée) et ce, tout au long de ce chapitre. Ceci nous permet, par exemple, de voir si la concentration des activités innovantes du début de la période d'étude (en 2002) est similaire ou bien différente de celle de la fin de la période (en 2008)

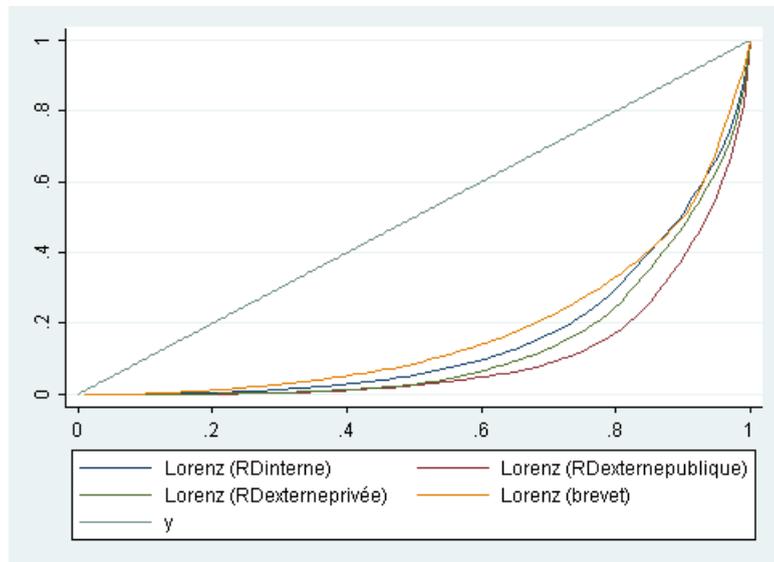


Figure 3.2 – Concentration des brevets et de la R&D interne et externe en 2008

La courbe de Lorenz montre clairement une forte concentration géographique de l'innovation pour les années 2002 et 2008. L'ampleur de la concentration est donnée par l'indice de Gini. Nous présentons dans le tableau suivant les valeurs de l'indice de Gini calculées pour chacune des variables R&D et brevets pour les années 2002 et 2008 :

Tableau 3.1 – Indice de Gini

	2002	2008	Rang
Brevet	0.600	0.637	4
$R\&D_{int}$	0.746	0.681	3
$R\&D_{extpub}$	0.851	0.777	1
$R\&D_{extpriv}$	0.840	0.726	2

A partir du tableau (3.1), nous remarquons que la R&D externe publique occupe la première place pour les deux années considérées avec une valeur de 0.851 en 2002 et 0.777 en 2008. Au deuxième rang, la R&D externe privée avec une valeur de 0.84 en 2002 et 0.726 en 2008. En troisième position se situe la R&D interne avec une valeur de 0.746 en 2002 et 0.681 en 2008. Les résultats montrent une forte concentration de la R&D et des brevets en France. Nos résultats sont similaires à ceux évoqués plus haut sur la concentration de l'innovation et ainsi qu'à d'autres travaux sur la concentration de l'emploi en France (Houdebine, 1999).

Ainsi, en s'intéressant à l'évolution de la concentration des activités en France entre 1978 et 1992⁵, Houdebine (1999) a conclu que la concentration géographique moyenne des activités a diminué pour cette période et que les départements français sont de moins en moins spécialisés. Ce résultat a été confirmé par Combes (2000a,b) pour une période allant de 1984-1993. Par ailleurs, les indicateurs utilisés pour mesurer la concentration présentent des limites (Arbia, 2001a ; Lafourcade et Mion, 2007). Par exemple, pour l'indice de Gini, la norme par rapport à laquelle nous raisonnons est la distribution uniforme ou celle de l'activité totale. Cependant, aucun test du caractère significatif de la déviation de l'indice de Gini par rapport à cette norme n'a jusqu'à présent été proposé (Combes et Duranton, 2006). Ainsi, des départements ou pays ayant le même indice de Gini peuvent présenter des distributions différentes d'innovation. D'ailleurs, la courbe de Lorenz peut avoir des formes différentes pour un même indice de Gini⁶.

Avec l'utilisation de ces indices de concentration a-spatiaux, nous ne pouvons donc pas avoir des informations sur la structure de localisation des activités innovantes. Autrement dit, nous ne pouvons pas savoir comment un département peut influencer l'innovation d'un autre et si ce sont les observations similaires ou celles dissemblables qui influencent le plus sur la concentration de l'innovation.

Il semble donc pertinent d'étudier la concentration qui prend en compte le positionnement de chaque observation, supposé non aléatoire. Ce type de concentration n'est autre que l'agglomération qui pose l'hypothèse de l'existence d'une autocorrélation spatiale entre les départements voisins.

3.2.2. En quoi la concentration diffère-t-elle de l'agglomération ?

Nous avons montré dans la sous section précédente que l'indice de Gini ainsi que les autres indices de concentration ne prennent pas en compte la position relative ni la distance qui sépare chaque département et qu'il n'est pas neutre quant à la mesure de l'agglomération⁷. La non considération de ces dimensions conduit à une mesure erronée de l'agglomération dans la mesure où la concentration ne tient pas compte des

5. Étude menée en utilisant les enquêtes structure des emplois (ESE) pour un découpage sectoriel en 100 niveaux de la nomenclature d'activité et de produit (NAF) et un zonage départemental.

6. Concernant l'indice d'Ellison et Glaeser, la référence n'est plus la distribution uniforme ou la distribution de l'emploi total, mais c'est la valeur prise par l'indice lorsque les établissements obéissent au principe d'une localisation aléatoire. En revanche, les seuils fixés pour mesurer le degré de la concentration géographique d'un secteur restent subjectifs. En outre, ces indices n'intègrent pas la position relative (*checkerboard problem*) ainsi que la proximité géographique des unités spatiales (White, 1983).

7. Le terme exact utilisé par Arbia (2001b) est "*polarisation*". Néanmoins, Lafourcade et Mion (2007) préfèrent utiliser le terme "*agglomération*" qui a été largement mobilisé par la NEG (Nouvelle Economie Géographique) afin de refléter réellement le processus de localisation résultant de

interactions spatiales entre les observations. Il serait donc intéressant d'apporter des informations complémentaires sur la structure spatiale des activités d'innovation et sur le processus de diffusion des connaissances.

Pour cela, nous exploitons les indices d'agglomération spatiale qui prennent en compte la localisation relative de nos départements (indice de Moran global et indice de Moran local). Ces indices vont nous permettre de déterminer si la répartition géographique des activités innovantes est aléatoire ou non. Les études récentes font de plus en plus appel à ces techniques afin d'apporter de nouveaux éclairages.

Le processus d'agglomération est un phénomène polymorphe qui nécessite d'explorer conjointement deux dimensions : d'une part, une mesure de la concentration de ces activités et, d'autre part, une caractérisation de la structure spatiale de cette localisation (Guillain et Le Gallo, 2007).

L'objectif principal de l'étude de l'agglomération des activités innovantes est de vérifier s'il existe une corrélation positive ou négative entre les activités d'innovation menées au sein des unités spatiales, géographiquement proches. Avant de présenter les outils utilisés pour mesurer l'agglomération des activités innovantes, nous illustrons dans ce qui suit la différence entre agglomération et concentration avec un exemple inspiré de l'étude faite par Arbia (2001b) :

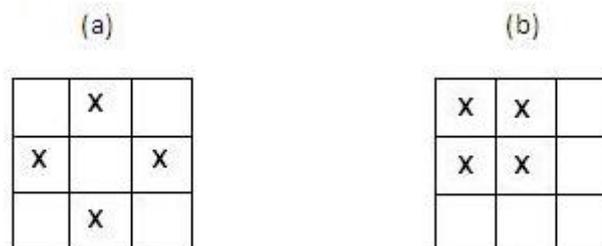


Figure 3.3 – Agglomération ou concentration ?

Source : Arbia, 2001b ; Lafourcarde et Mion, 2007

Il s'agit de répartir 4 usines dans une grille de 3×3 cellules. Dans la figure 3, l'activité est inégalement répartie de deux façons différentes :

- la case *a* illustre la concentration, un concept de distribution qui n'est pas affecté par la permutation des observations dans l'espace. Autrement dit, la position relative de chaque observation dans l'espace n'est pas prise en compte dans ce premier cas. La valeur des indices de concentration présentés dans la sous section précédente (tel que l'indice de Gini) est la même dans le cas *a* et *b*. En effet,

l'interaction entre les coûts de transport (et donc la distance) et les rendements d'échelle croissants. Ceci a été expliqué en détail dans la section 3 du chapitre 1 de notre travail.

ces indices permettent de quantifier la variabilité spatiale de chaque observation par rapport à la moyenne, sans considérer la distance entre les différentes unités spatiales.

- En revanche, la case *b* reflète l’agglomération parce qu’elle témoigne d’un certain degré de dépendance spatiale, qui se traduit dans le coin gauche de la case *b*. Plus précisément, nous parlons d’autocorrélation spatiale si, chaque usine d’une industrie donnée et d’une région *i*, est informée d’une manière directe de la localisation des autres usines de la même industrie des régions voisines (Lafourcarde et Mion, 2007). Cette relation linéaire entre les différentes unités spatiales est plus manifestée dans la case (b) que dans la case (a).

Afin de mesurer l’agglomération, nous mobiliserons les outils de l’analyse exploratoire de données spatiales (AEDS)⁸ que nous présenterons dans la section 4 de ce chapitre. Contrairement aux outils de mesure de la concentration dont le résultat est invariant à une permutation, l’AEDS permet la prise en compte de la position relative des unités au sein de la zone étudiée. Avant de présenter ces outils, nous montrons maintenant l’importance de la prise en compte de la dépendance spatiale.

3.3. Intérêt d’une prise en compte de la dépendance spatiale

Deux raisons principales peuvent être attribuées au regain de l’attention portée à la prise en compte de l’autocorrélation spatiale entre observations. La première est l’importance de la prise en compte de l’effet de la proximité géographique à travers une méthode plus directe que la FPI. En effet, la modélisation de l’autocorrélation spatiale permet de capter l’existence, l’ampleur et l’influence des effets de débordements géographiques⁹ (Le Gallo, 2002). La deuxième raison, d’ordre exploratoire, consiste à étudier comment les observations sont liées entre elles, tout en montrant l’influence effective (positive ou négative) qu’exerce une localisation particulière par rapport à d’autres.

Afin d’étudier les interactions spatiales, il est nécessaire de présenter les différents types d’autocorrélation spatiale ainsi que les liens de voisinage qui existent entre chaque élément de l’espace étudié.

8. Exploratory Spatial Data Analysis ou ESDA.

9. Objectif du chapitre 4 de notre thèse.

3.3.1. Définition et sources de l'autocorrélation spatiale

La présence de l'autocorrélation spatiale suppose que les valeurs prises par une variable aléatoire dans un espace géographique ne sont pas disposées au hasard, mais sont souvent semblables pour deux observations spatiales proches ou contiguës (Jayet, 1993).

Anselin et Bera (1998) définissent l'autocorrélation spatiale comme étant la coïncidence entre *similitude des attributs* et *similitude des localisations* : "*Spatial autocorrelation can be loosely defined as the coincidence of value similarity with locational similarity*". En d'autres termes, c'est la corrélation, positive ou négative, d'une variable avec elle-même provenant de la disposition géographique des données :

- L'autocorrélation spatiale positive se traduit par une tendance à la concentration dans l'espace d'observations de classes voisines : les lieux proches se ressemblent davantage que les lieux éloignés.
- L'autocorrélation spatiale négative se traduit par le regroupement géographique d'observations très différentes : des lieux proches sont plus différents que des lieux éloignés.
- Enfin, l'absence d'autocorrélation spatiale indique que la répartition spatiale des valeurs de la variable est aléatoire : aucune relation n'existe entre la proximité des lieux et leur degré de ressemblance (Griffith, 1992b)

En suggérant la première loi de la géographie suivante, "*Everything is related to everything else, but closer things more so*", Tobler (1979) avait déjà souligné la présence d'autocorrélation spatiale entre ce qui se passe en un point de l'espace et ce qui se passe ailleurs. En effet, détecter de l'autocorrélation spatiale dans une variable fournit 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 ; Laurent et al., 2012).

Or, il est intéressant de noter qu'en économie, nous sommes généralement confrontés à trois types de localisation. Il peut s'agir, tout d'abord, de points représentant, par exemple, des localisations d'unités de production ou de distribution (mesurés par leur latitude et leur longitude). Ensuite, il peut s'agir de lignes, connectées entre elles ou non, comme un réseau routier ou fluvial. Enfin, ces localisations peuvent être des zones géographiques comme des régions ou des pays. Dans les trois cas, le nombre de ces points, de ces lignes ou de ces zones est supposé fini (Le Gallo, 2002).

Par ailleurs, l'autocorrélation spatiale présente quelques particularités par rapport à la dimension temporelle. Cette dernière est unidimensionnelle puisque seul le passé influence le futur. Contrairement à l'autocorrélation temporelle, l'autocorrélation

spatiale est multidimensionnelle puisque "*tout est relié à tout*" (Zaninetti, 2005). L'autocorrélation spatiale a deux sources principales :

- 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 mouvements de biens, de personnes et de capitaux en un lieu donné affectent les conditions en d'autres lieux, ce qui constitue une source d'autocorrélation spatiale. Ainsi, la diffusion d'un phénomène 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.
- 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.

Après avoir décrit la manière dont sont présentées les observations dans l'espace, nous présentons dans ce qui suit comment ces interactions sont d'autant plus importantes que les localisations en question sont "*proches*" les unes des autres. Dès lors, nous avons besoin d'un instrument afin de modéliser ces interactions entre observations et leur décroissance en fonction de la distance qui les sépare. Cet instrument est appelé matrice d'interactions spatiales ou matrice de poids.

3.3.2. Les matrices de poids

Un des éléments fondamentaux de la mesure de la structure spatiale est le choix de la matrice de poids qui vient contraindre l'étendue de la dépendance spatiale. Le choix des matrices de poids se révèle crucial puisqu'il conditionne très fortement les résultats que nous allons obtenir ainsi que l'analyse qui en découle.

Dans notre cas, la matrice de poids que nous utiliserons dans le calcul des indices de Moran et dans l'estimation de nos modèles économétriques spatiaux se fonde sur les caractéristiques géographiques des unités spatiales retenues, à savoir les 94 départements français. Il s'agit, en effet, de définir un ensemble de voisins pour chaque département, ce qui aboutit aux matrices de poids. 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 par définition une matrice carrée, ayant autant de lignes et de colonnes qu'il y a de zones géographiques avec 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. Trois catégories de matrice de voisinage sont généralement utilisées dans l'économétrie spatiale : les matrices de contiguïté, les matrices des *k-plus proches voisins* et les matrices de distance.

3.3.2.1. Matrices de contiguïté

Ces matrices sont les plus utilisées en raison de leur simplicité. La contiguïté entre deux régions se définit par le fait qu'elles ont une frontière commune et que chaque terme de cette matrice est égal à 1 si les régions sont contiguës à l'ordre 1 et 0 sinon. Par convention, une région n'est pas contiguë avec elle-même : $W_{ii} = 0, \forall i$.

Pour 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, $L_i = \sum_{j=1}^N W_{ij}$.

Le nombre total de liens existant dans le système régional est alors égal à

$$A = \frac{1}{2} \sum_{j=1}^N L_i.$$

Cette définition entraîne que la matrice W est symétrique. La notion de contiguïté peut être généralisée : 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 .¹⁰

Cliff et Ord (1973) ont défini la contiguïté de trois manières différentes :

- Soit un seul point de frontière commune suffit, que nous appelons "*le cas du fou*",
- Soit deux points de frontière commune, donc un segment commun est nécessaire, "*le cas de la tour*",
- Soit les deux situations précédentes "*le cas de la reine*".

Nous représentons dans la figure (3.4) les différentes formes de contiguïté :

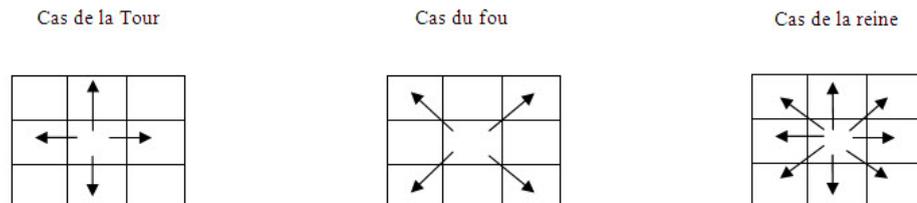


Figure 3.4 – Différentes formes de contiguïté

Source : Cliff et Ord (1973)

Généralement, les logiciels d'analyse spatiale utilisent le "*cas de la reine*" qui représente une définition plus large de la contiguïté. Le choix du type de contiguïté a une influence majeure sur la statistique calculée puisque *le cas de la tour* ne donne que quatre cellules voisines tandis que celui de *la reine* donne huit cellules voisines à chaque quadrant.

10. La matrice de contiguïté à l'ordre k n'est pas égale à la matrice de contiguïté à l'ordre 1 élevée à la puissance k . Blommestein et Koper (1992) et Anselin et Smirnov (1996) ont développé des algorithmes pour pouvoir passer d'une matrice de contiguïté d'ordre 1 à une matrice d'ordre quelconque.

3.3.2.2. Les matrices des *k-plus proches voisins*

La forme la plus générale d'une matrice de *k-plus proches voisins* repose sur la distance entre les centroïdes géographiques des unités. La présentation et l'utilisation de ce type de matrice sont assez récentes (Pace et LeSage, 2004 ; Baumont et al., Ertur et le Gallo, 2004). En faisant appel à ce type de matrice, nous considérons que chaque unité est connectée au même nombre de voisins, peu importe sa localisation. Sa forme générale est définie de la manière suivante :

$$W_{ij} = \begin{cases} 0 & \text{si } i = j, \forall k \\ 1 & \text{si } d_{ij} \leq d_{ij}(k); \\ 0 & \text{si } d_{ij} > d_{ij}(k). \end{cases}$$

$W_{ij}(k)$ est un élément de la matrice de connectivité spatiale et $d_{ij}(k)$ est la plus petite distance d'ordre k entre les unités i et j , telle que l'unité i possède exactement k unités voisines¹¹.

3.3.2.3. Les matrices de distance

Cette matrice est construite à partir de la distance qui sépare deux unités géographiques i et j telles que les pays, les régions, les départements ou les entreprises. Cette distance peut être mesurée en utilisant les centroïdes des régions ou des départements, les capitales pour le cas des pays, ou bien dans le cas d'utilisation des observations réduite à un point. L'unité de mesure servant à définir la distance peut être par exemple la distance à vol d'oiseau, la distance parcourue par route ou le temps de transport (heures de vol par exemple ou heures de voiture ou de train...). Un élément de la matrice de poids peut s'écrire comme suit :

$$W_{ij} = (d_{ij})^{-\alpha} (\beta_{ij})^b.$$

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 et a et b sont des paramètres exogènes déterminés a priori.

Diverses formes sont disponibles :

– La fonction exponentielle négative $W_{ij} = e^{-ad_{ij}}$

– Ou une fonction de l'inverse de la distance $W_{ij} = \begin{cases} d_{ij}^{-\beta} & \text{si } d_{ij} < \tilde{d} \\ 0 & \text{sinon} \end{cases}$

11. Voir annexe 2 pour plus de détail sur la construction des matrices des *k-plus proches voisins*. Dans cet annexe, nous avons représenté graphiquement les voisins d'ordre 1 à 4 du département Haute-Garonne.

Ou α et β sont des paramètres déterminés a priori, \tilde{d} est la valeur seuil au delà de laquelle nous supposons que les régions i et j ne sont pas connectées.

3.3.2.4. Standardisation des matrices de poids

La standardisation en ligne d'une matrice de poids est une opération qui consiste à rendre la somme des poids des voisins égale à 1 pour chaque unité. L'objectif principal de la standardisation est de faire une correction nécessaire afin que chaque unité ait le même poids quelque soit le nombre de voisins¹².

Chaque élément de la matrice W_{ij} de la matrice est divisé par la somme totale de la ligne :

$$W_{ij}^s = \frac{W_{ij}}{\sum_j W_{ij}}$$

Ces matrices standardisées ne sont plus nécessairement symétriques. Les poids sont alors compris entre 0 et 1 et cette opération permet de comparer les paramètres spatiaux dans différents modèles économétriques.

Par ailleurs, la standardisation des matrices de poids change l'interprétation : par exemple, dans le cas d'une matrice de distance, la connexion entre deux régions après standardisation ne dépend plus de la distance absolue mais de la distance relative entre ces régions. L'utilisation des matrices de poids conduit à la possibilité de relier une observation de la variable en un point de l'espace avec les observations de cette variable pour d'autres lieux du système. C'est ce que nous appelons les opérateurs spatiaux ou bien les variables spatialement décalées : "*Spatial Lag*" (Le Gallo, 2002). Nous définissons la variable spatialement décalée (*Spatial Lag*) pour N régions par $WX = W \times X$. Il s'agit d'un vecteur X avec une valeur x pour chaque région i et d'une matrice de poids W . Lorsque W est une matrice standardisée, l' $i^{\text{ème}}$ élément de la variable spatialement décalée contient la moyenne pondérée des observations des régions voisines à la région i . Les décalages spatiaux d'ordre supérieur sont définis par W_x^2, W_x^3 , 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, nous parlons d'autocorrélation spatiale positive entre Y_i et ses voisines alors que si Y_i et $(WY)_i$ sont différents, il y a autocorrélation spatiale négative entre Y_i et ses voisins¹³.

Après avoir étudié les différentes matrices de poids et leurs caractéristiques, nous allons présenter dans ce qui suit les matrices que nous adopterons dans notre analyse globale et locale de l'autocorrélation spatiale entre départements.

12. À défaut, les unités ayant le plus de voisins auront plus de poids dans le calcul de l'autocorrélation spatiale. Ceci reste valable aussi bien à l'échelle globale que locale

13. Dans le diagramme de Moran présenté dans la section 4 de ce chapitre 3, la variable *Spatial Lag* est représentée par l'ordonnée du graphique.

3.3.3. Comparaison et critère de choix des matrices

Les matrices de poids sont d'une importance fondamentale pour introduire l'autocorrélation spatiale dans les modèles économétriques. En effet, pour modéliser les interactions spatiales, il est nécessaire d'imposer une structure sur l'étendue de ces interactions. Pour ce faire, nous construisons la matrice de contiguïté, la matrice de distance de moins de 100km, la matrice de distance de moins de 200km¹⁴ et la matrice des *k-plus proches voisins* :

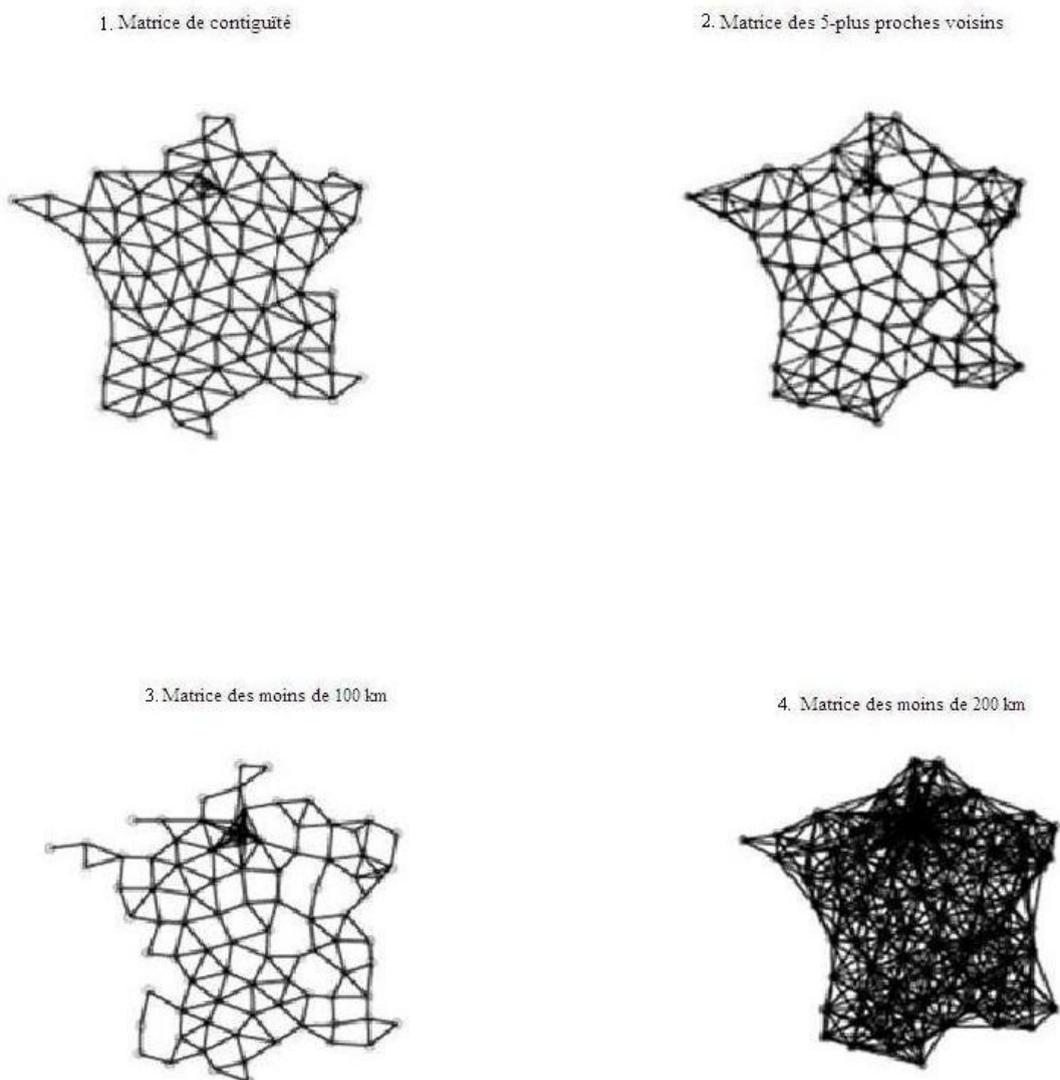


Figure 3.5 – Matrices de poids utilisées

14. L'objectif est de pouvoir comparer nos résultats avec ceux trouvés en utilisant des découpages géographiques par zone de 200 kilomètres et de 100 kilomètres.

La structure spatiale de nos données est déterminante dans le choix de la matrice de poids appropriée. Une analyse de la structure spatiale de nos données montre que le nombre moyen des voisins à moins de 200 kilomètres (respectivement à moins de 100 kilomètres) est de l'ordre de 23,255 départements (respectivement 11,627). Étant donné la forme géographique et la petite taille des départements français, l'utilisation des matrices des voisins à moins de 200 km n'est sans doute pas la technique la plus appropriée pour notre échantillon. D'ailleurs, la carte n°4 de la figure (3.5) illustre bien le nombre élevé attribué à chaque département notamment aux départements de la région Île-de-France. A l'inverse, en utilisant une matrice de distance de moins de 100km, nous sommes confrontés à un problème d'effet de bord. En effet, 82 départements se sont retrouvés sans voisins (voir carte 3 de la figure 3.5). Pour toutes ces raisons, nous choisissons d'utiliser dans nos estimations deux types de matrices de poids : la matrice de contiguïté d'ordre 1 et la matrice des *k*-plus proches voisins¹⁵. Ces matrices sont les plus utilisées (Fingleton et al., 1997 ; Fingleton, 1999, 2000 ; Le Gallo, 2002).

Le choix de la première est justifié par le fait qu'elle tient compte des "*voisins réels*" à chaque département. Nous savons par exemple que les Haute-Pyrénées, le Gers, le Tarn et Garonne, le Tarn et l'Ariège ont des frontières communes avec le département Hautes-Garonne. L'utilisation d'une matrice de contiguïté d'ordre 1¹⁶ dans notre cas, permet d'étudier l'effet de ces départements voisins (ayant une frontière commune) sur l'innovation du département Hautes-Garonne par exemple. Par contre, en utilisant la matrice de distance de moins de 200km, d'autres départements seront pris en compte dans la mesure de contiguïté (comme Le Lot et l'Aveyron) même s'ils ne constituent pas réellement des voisins au département considéré.

La deuxième matrice de poids (des *k*-plus proches voisins) a été adoptée par plusieurs auteurs (Baumant et al, 2003 ; Guillain et Le Gallo, 2007 ; Autant-Bernard et Le Sage, 2011). Comme nous l'avons évoqué plus haut, les matrices des *k*-plus proches voisins attribuent le même nombre de voisins à chaque département. Ce nombre noté *k*, représente le nombre moyen de voisins par département :

Tableau 3.2 – Répartition du nombre de voisins par département

Nombre de voisins	2	3	4	5	6	7	8	10
Nombre de départements	6	12	14	19	30	11	1	1

15. Pour simplifier, l'analyse exploratoire et les estimations du chapitre 4 ont été réalisées avec la matrice de contiguïté d'ordre 1. Nous avons aussi refait les mêmes estimations et construit les mêmes cartes avec la matrice des 5-plus proches voisins. Les résultats trouvés sont très proches.

16. C'est le cas de la reine qui est utilisé avec le logiciel *R* afin de construire notre matrice de contiguïté.

Le département avec le plus de voisins est la Seine-et-Marne (77) avec 10 voisins. Les départements avec le moins de voisins sont les Alpes-Maritimes (06), le Finistère (29), la Moselle (57), le Pas-de-Calais (62), les Pyrénées-Orientales (66) et la Haute-Savoie (74) avec 2 voisins.

Pour déterminer la valeur de k , nous construisons le tableau suivant :

Tableau 3.3 – Nombre total et moyen de voisins par département

Nombre de départements	Nombre total de liens ou de voisins entre départements	Nombre moyen de voisins
94	474	5.042

Le nombre k de voisins les plus proches est donc fixé à 5. Nous choisissons la matrice de contiguïté et la matrice des 5-plus proches voisins pour modéliser les interactions spatiales entre départements français¹⁷.

Nous présentons dans la section suivante les méthodes d'analyse exploratoire des données spatiales qui permettent de déterminer la présence et l'ampleur de l'effet de l'autocorrélation spatiale formalisée avec la notion de matrices de poids.

3.4. Analyse exploratoire de données spatiales et identification des *clusters* innovants en France

Deux approches sont de plus en plus considérées pour analyser la dépendance spatiale entre observations. La première possibilité est offerte par l'analyse exploratoire de données spatiales (objet de cette section). Quant à la deuxième possibilité, elle est réalisée grâce aux outils de l'économétrie spatiale (objet du chapitre 4).

Nous mobilisons donc dans cette section les outils de l'analyse exploratoire de données spatiales (AEDES)(Getis et Ord, 1992; Anselin, 1995; Anselin et Bera, 1998; Le Gallo, 2002) afin de fournir des éléments de réponse à la question de la répartition spatiale des brevets et de la R&D, ainsi qu'à la mesure de la dynamique d'agglomération. L'évaluation des interactions spatiales entre régions avec les outils AEDES permet de décrire et de visualiser les distributions spatiales, d'identifier les localisations atypiques et les points extrêmes, de détecter les schémas d'association spatiale et enfin de suggérer des régions spatiales ou d'autres formes d'hétérogénéité spatiale (Anselin, 1998a,b et 1999; Bailey et Gatrell, 1995; Ertur et Koch, 2007).

L'analyse avec les outils AEDES permet l'exploration de deux dimensions : la mesure de l'autocorrélation spatiale globale d'une part et, d'autre part, une caractérisation de

17. Le Gallo (2002) montre que les résultats obtenus restent similaires quelque soit les matrices de poids utilisées (matrice de contiguïté ou matrice des k-plus proches voisins).

la structure locale de cette autocorrélation (Guillain et le Gallo, 2007). Nous résumons les outils AEDES en deux groupes comme le montre le schéma suivant :

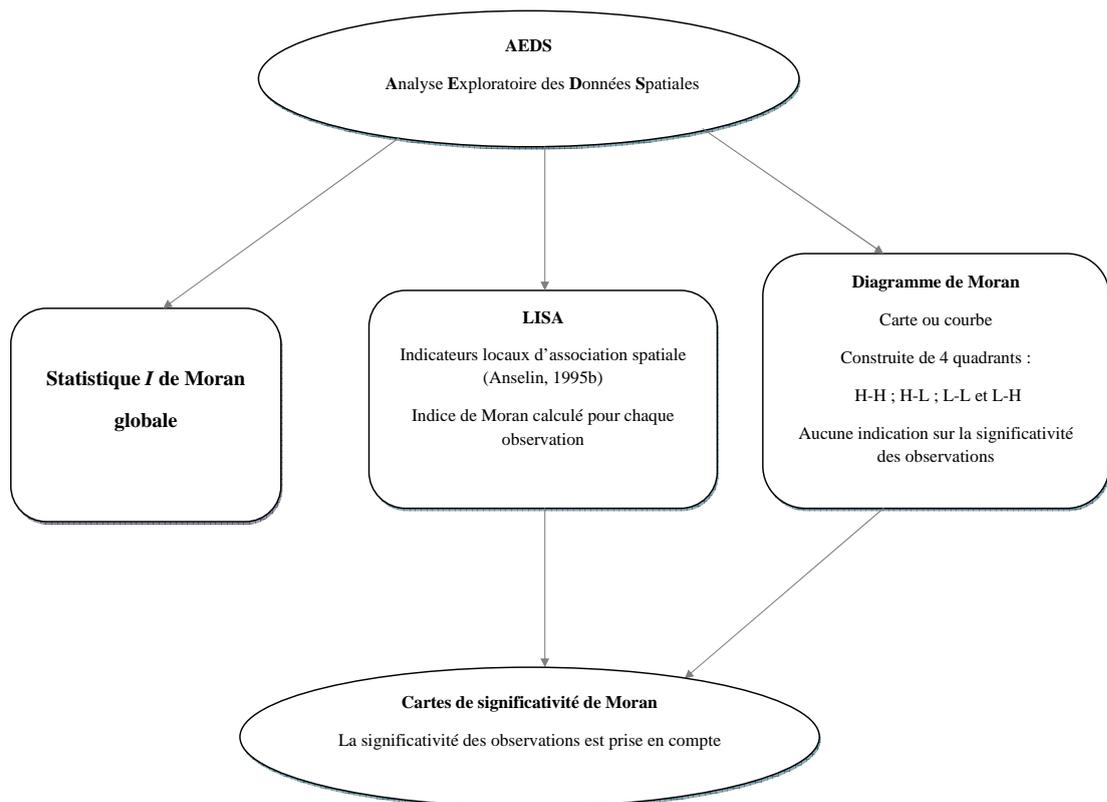


Figure 3.6 – Les outils de l’analyse exploratoire de données spatiales (AEDES)

Comme le montre la figure (3.6), ces outils sont répartis en un outil global : statistique de Moran globale, et des outils d’analyse locale : diagramme de Moran, statistique locale de Moran I ou statistiques de LISA et cartes de significativité de Moran.

En effet, la statistique de Moran est une statistique globale qui ne permet pas d’apprécier la structure régionale de l’autocorrélation spatiale. Une version locale de cet indicateur permet d’affiner cette analyse globale afin de caractériser le phénomène d’autocorrélation spatiale au niveau des entités géographiques étudiées. Nous utiliserons donc l’indicateur d’association spatiale (LISA)¹⁸ pour cette fin.

L’utilisation combinée du diagramme de Moran et de la statistique de LISA permet d’obtenir ce que nous appelons les cartes de significativité de Moran. Ces cartes permettent en même temps d’identifier les quatre types d’autocorrélation spatiale locale et de connaître si cette autocorrélation observée est significative. En utilisant

18. Local Indicators of Spatial Association.

cette approche, nous aurons plus d'évidences empiriques sur le caractère spatial de la connaissance par rapport à la simple observation de l'autocorrélation.

3.4.1. Analyse globale de l'autocorrélation

Pour évaluer l'importance de l'agglomération, nous observons l'existence de l'autocorrélation spatiale globale entre les observations au niveau des départements considérés. Il est néanmoins nécessaire de choisir l'indice utilisé (parmi les plus connus : la semi-variance, l'indice de Moran, le coefficient de Geary). Notre choix s'est porté sur l'indice de Moran (I de Moran).

3.4.1.1. Statistique I de Moran globale

Il existe trois indices de mesure du degré de similarité des réalisations voisines des variables spatialisées : un pour les variables qualitatives (le test des couleurs ou le " *joint count statistic*") et deux pour les variables quantitatives (Indice de Moran 1948, 1950 a et b et indice de Geary, 1954). Comme nous disposons de variables quantitatives, nous nous limiterons à l'étude des deux derniers indices.

Encadré 3.1 : Indice d'autocorrélation spatiale des variables qualitatives

Le test des couleurs consiste à superposer une grille de carroyage à l'espace d'étude. Nous créons un caractère qualitatif binomial indiquant la présence (codée 1) ou l'absence (codée 0) d'un point du semis dans chaque cellule du carroyage. Sur une carte, nous pourrions marquer en noir (n) chaque cellule où le caractère est présent, et laisser en blanc (b) les cellules où le caractère est absent, ce qui produit une sorte de damier.

Selon le type de contiguïté choisi, nous pouvons joindre les cellules adjacentes selon leurs couleurs : "*nn*" pour dire que le caractère est présent dans deux unités voisines, "*bb*" lorsque les deux unités voisines sont vides et "*nb*" si le caractère est présent dans l'une des unités territoriales et pas dans l'unité voisine. Nous parlons donc d'autocorrélation spatiale positive lorsque nous dénombrons un nombre de contacts "*nn*" ou "*bb*" plus grand que les contacts "*nb*".

L'autocorrélation spatiale est négative lorsque nous obtenons un nombre maximum de contacts "*nb*" et relativement peu de "*bb*" et de "*nn*".

Ainsi, nous pourrions calculer les tests NB et NN comme suit :

- Le test NB consiste à comparer le nombre de paires nb observées aux paramètres d'une distribution de probabilités $P(nb)$ dans l'hypothèse nulle où le caractère est reparti aléatoirement selon une loi binomiale :

$$\epsilon_{NB} = \frac{NB - \mu_{NB}}{\sqrt{V(NB)}}$$

- Si $|\epsilon_{NB}| \geq 1.96$, ceci signifie qu'il existe une autocorrélation spatiale significative, que les événements ont tendance, soit à se localiser à côté les uns des autres, ce qui se traduit par un nombre de paires "nb" observé sensiblement plus faible que l'espérance théorique d'un processus ponctuel stationnaire, soit à s'éloigner les uns des autres, ce qui se traduit par un nombre de paires nb observé sensiblement plus élevé que l'espérance théorique d'un processus ponctuel stationnaire.
- Le test NN consiste à comparer le nombre de paires "nn" observée aux paramètres d'une distribution de probabilités $P(nn)$ dans l'hypothèse nulle où le caractère est réparti aléatoirement selon une loi binomiale.

$$\epsilon_{NN} = \frac{NN - \mu_{NN}}{\sqrt{V(NN)}}$$

- Le test NN est soumis aux mêmes conditions d'application que le test NB.
- L'indice de Moran (respectivement l'indice de Geary) est la généralisation du test NN (respectivement NB) aux caractères quantitatifs.

Au début, la littérature en économétrie spatiale était dominée par *le test du ratio de vraisemblance* et *le test de Wald* (Brandsma et Kelletaper, 1979; Anselin, 1980; Cliff et Ord, 1981).

A l'heure actuelle, les statistiques de l'autocorrélation spatiale les plus utilisées sont la statistique de Moran (I) et la statistique de Geary (C) qui a une fonction voisine.

La statistique I de Moran s'écrit sous cette forme :

$$I = \frac{\sum_i \sum_j W_{ij} (x_i - \bar{x})(x_j - \bar{x}) / S_0}{\sum_i (x_i - \bar{x})^2 / N} \quad (3.1)$$

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, W_{ij} est l'élément de la matrice de poids et 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'indice de Moran s'interprète comme un coefficient de corrélation classique. Généralement, il varie entre -1 (*autocorrélation spatiale négative indiquant la concentration des valeurs dissimilaires*, Anselin, 1995) et 1 (*autocorrélation spatiale positive indiquant une concentration des valeurs similaires*).

La statistique de Geary (1954) s'écrit de la façon suivante :

$$C = \frac{(N - 1) \sum_i \sum_j W_{ij} (x_i - \bar{x})^2}{2S_0 \sum_i (x_i - \bar{x})^2} \quad (3.2)$$

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.

Ces deux indices permettent de valider l'existence d'autocorrélation spatiale pour une variable donnée, et ce pour l'ensemble des points. L'utilisation de la statistique de Moran ou celle de Geary donne des résultats voisins. D'un point de vue statistique, le I de Moran semble conserver un avantage sur le C de Geary (Cliff et Ord, 1981) pour les raisons suivantes :

- Le I de Moran est moins fragile que le second car il mesure les écarts à la moyenne et est donc peu sensible aux valeurs aberrantes.
- La similarité de son écriture et celle de ses résultats avec le coefficient de corrélation le rendent plus facile d'accès.
- Finalement, l'indice de Moran est très répandu dans les logiciels géostatistiques, statistiques, voire dans les systèmes d'information géographiques (SIG)(Oliveau, 2005). Ceci constitue un argument d'ordre pragmatique expliquant l'utilisation importante de cet indice dans la mesure de l'autocorrélation spatiale.

Pour toutes ces raisons, notre choix s'est porté sur *l'indice de Moran*¹⁹.

Par ailleurs, il est nécessaire de définir la matrice de voisinage pour calculer le I de Moran. Compte tenu de la configuration spatiale des départements caractérisée par une hétérogénéité au niveau des tailles des départements, nous retenons la matrice de contiguïté pour calculer le I de Moran. Dans le tableau suivant, les résultats sont obtenus en utilisant une matrice de contiguïté d'ordre 1²⁰. Nous mesurons donc l'agglomération en utilisant le I de Moran global calculé pour chaque variable :

19. Comme pour les statistiques NN et NB , le test de Moran est calculé en utilisant l'espérance de I de Moran, $E(I)$ et sa variance, $V(I)$. La statistique de I de Moran noté Z s'écrit comme suit :

$Z = \frac{I - E(I)}{\sqrt{V(I)}}$. La statistique de ce test est asymptotiquement normale sous l'hypothèse nulle

d'absence d'autocorrélation spatiale. Si Z est significativement positive (respectivement négative), nous concluons à l'existence d'autocorrélation spatiale positive (respectivement négative). Ce test est généralement utilisé après régression afin de détecter l'autocorrélation spatiale au niveau des résidus.

20. les valeurs de I de Moran calculées pour une matrice des 5-plus proches voisins conduisent aux mêmes conclusions quant au signe et à la significativité de l'autocorrélation spatiale globale.

Tableau 3.4 – Indice de Moran

	2002	2008	Rang
Brevet	0.196 (0.000)	0.287 (0.000)	1
$R\&D_{int}$	0.195 (0.000)	0.242 (0.000)	2
$R\&D_{extpub}$	0.060 (0.053)	0.127 (0.000)	4
$R\&D_{extpriv}$	0.145 (0.000)	0.234 (0.000)	3

Les valeurs du I de Moran représentées ci-dessous sont obtenues avec le logiciel statistique R ²¹. Ces résultats montrent une faible agglomération des activités innovantes. La variable la plus spatialement autocorrélée correspond aux brevets. Cette variable occupe la première position et ce, pour les années 2002 et 2008 avec une valeur de 0.196 en 2002 (respectivement 0.287 en 2008). En deuxième position nous trouvons la R&D interne, ensuite la R&D externe publique puis en dernier lieu la R&D externe privée. Toutes ces variables sont considérées comme agglomérées au sens de la statistique I de Moran global à un seuil de 1% sauf l'indice de Moran de la R&D externe publique qui est significatif seulement à un seuil de 10%. Nous remarquons cependant en comparant avec l'indice de Gini, que ces variables fortement concentrées ne sont pas nécessairement fortement autocorrélées spatialement.

Toutefois, la statistique de Moran considérée constitue une mesure globale de l'autocorrélation spatiale. Elle ne nous permet donc pas d'analyser la structure locale de cette autocorrélation. Pour cela, nous exploitons les outils de l'autocorrélation spatiale locale pour déterminer quels départements contribuent majoritairement à l'autocorrélation spatiale globale. Autrement dit, nous cherchons où se situent les regroupements de valeurs fortes (respectivement faibles). Les outils utilisés sont le diagramme de Moran et les statistiques de LISA. Utiliser conjointement les analyses du diagramme de Moran et la significativité des LISA nous permet de savoir, d'une part, dans quel quadrant se situe l'unité spatiale i (département pour notre cas) et, d'autre part, si la situation de l'unité spatiale i est significativement influencée par celles de ses voisines et réciproquement (Chantelot et al, 2010).

21. Même résultats sont obtenus avec le logiciel GeoDa en utilisant une procédure de permutation d'Anselin (1995), avec 999 permutations (p-value=0.001), dans laquelle une distribution de référence est générée empiriquement pour le I de Moran. Nous obtenons alors des niveaux de significativité qui permettent de conclure le test.

Nous présentons dans la sous section suivante les diagrammes de Moran des dépenses en R&D et des brevets afin de détecter les regroupements spatiaux (*clusters*) et d'analyser l'instabilité départementale représentée sous la forme de localisations atypiques.

3.4.2. Analyse locale de l'autocorrélation

Après avoir détecté l'existence d'une autocorrélation spatiale globale au niveau des brevets et des dépenses en R&D pour les années 2002 et 2008, nous nous focalisons dans ce qui suit sur l'autocorrélation spatiale locale entre départements. Une analyse par les outils de l'autocorrélation spatiale locale permet d'affiner cette analyse globale tout en permettant d'apprécier la structure départementale ainsi que de caractériser la nature du phénomène d'autocorrélation spatial au niveau de ces départements étudiés.

L'autocorrélation spatiale locale diffère de celle globale sur les points suivants :

Tableau 3.5 – Différences entre méthodes globales et méthodes locales de l'autocorrélation spatiale

Analyse globale	Analyse locale
Résume l'information pour la région entière	Désagrégation locale des méthodes globales
Calcule un indice unique	Calcule des indices multiples
Résultat numérique qui ne se représente pas sur une carte	Résultat qui peut être représenté sur une carte
Méthode a-spatiale ou implicitement spatiale	Méthode explicitement spatiale
Met l'accent sur les similarités entre les territoires	Met l'accent sur les différences entre les territoires
Recherche des régularités ou des "lois"	Recherche des singularités ou des centres d'intérêt locaux (hot spots)
<i>Exemple</i> : méthode du voisin le plus proche	<i>Exemple</i> : indices locaux de dépendance spatiale (LISA)

Source : Fotheringham et al.(2002)

Après avoir présenté la différence entre une analyse globale et une analyse locale de l'autocorrélation spatiale, nous nous intéresserons dans cette section aux outils utilisés pour établir ces analyses. L'objectif sera donc double : nous souhaiterons dans un premier temps, savoir s'il existe des regroupements locaux de valeurs fortes ou faibles (diagramme de Moran), et de déterminer, dans un deuxième temps, quels sont les départements qui contribuent le plus à l'autocorrélation spatiale globale tout en déterminant les départements qui s'éloignent du schéma global d'autocorrélation spatiale (statistiques de LISA).

3.4.2.1. Diagramme de Moran

L'instabilité spatiale locale est analysée à l'aide du diagramme de Moran (Anselin, 1996). Ce diagramme permet de visualiser les formes de l'autocorrélation spatiale locale en visualisant l'association linéaire entre X et WX sous la forme d'un graphique.

Il est construit sur la base des valeurs centrées réduites : l'abscisse représente la valeur centrée réduite du département considéré $X_i^* = X_i - E(X)$, et l'ordonnée indique la valeur moyenne des voisins de ce département. Cette moyenne pondérée (par les poids spatiaux) des observations associées aux localisations voisines est aussi appelée "variable décalée".

Elle représente ainsi l'information relative au voisinage de chaque localisation. Ainsi, l'ordonnée du point X_i^* peut s'écrire comme $\sum W_{ij} X_j^*$:

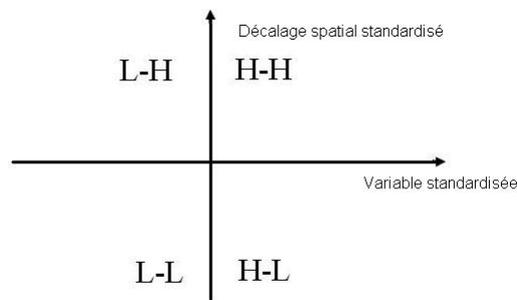


Figure 3.7 – Le diagramme de Moran

Le diagramme de Moran permet de répartir les observations sur quatre quadrants :

- *H-H* : Le quadrant "high-high" des valeurs autocorrélées positivement et supérieures à la moyenne : un département à valeur élevée (ou haute) entourée de départements à valeurs élevées.
- *L-L* : Le quadrant "low-low" des valeurs autocorrélées positivement et inférieures à la moyenne : un département à valeur faible (ou basse) entourée de départements à valeurs faibles.
- *H-L* : Le quadrant "high-low" des valeurs autocorrélées négativement où les observations représentées sont supérieures à la moyenne, tandis que la moyenne de leurs voisins est inférieure à la moyenne : un département à valeur élevée (ou haute) entourée de départements à valeurs faibles.
- *L-H* : Le quadrant "low-high" des valeurs autocorrélées négativement où les observations représentées sont inférieures à la moyenne, tandis que la moyenne de leurs voisins est supérieure à la moyenne : un département à valeur faible (ou basse) entourée de départements à valeurs élevées.

Pour notre cas, *l'axe horizontal* du diagramme de Moran représente le nombre de brevets (ou bien les dépenses en R&D) par département et *l'axe des ordonnées* mesure la moyenne du nombre de brevets des voisins (ou bien dépenses en R&D)²². L'intérêt d'exprimer les variables en forme standardisée (c'est-à-dire avec une moyenne nulle et un écart-type égal à 1) est d'estimer à la fois l'association spatiale globale (puisque la pente de la droite à l'intérieur du diagramme de Moran est équivalente au I de Moran) et l'association spatiale locale, à l'aide du diagramme de Moran dans lequel se situe le département. Nous pouvons ainsi mettre en évidence des *localisations atypiques* ou des *poches de non stationnarité locale* : un département innovant (de type H-L dans la carte) au milieu de départements non innovants appelé "*îlot de richesse*" ou à l'inverse un département non innovant (de type L-H) au milieu de départements riches "*moutons noirs*". Une des limites de cet outil est qu'il ne fournit pas d'information sur la significativité des regroupements spatiaux. Celle-ci peut par contre être obtenue à l'aide des indicateurs locaux d'associations spatiales (LISA).

L'utilisation des indicateurs LISA, développés par Luc Anselin au milieu des années 1990, semble prometteuse (Anselin, 1995). En effet, les LISA permettent le passage d'un indicateur global de la structure spatiale à une mesure locale des ressemblances. Ils constituent des indicateurs satisfaisant deux critères. Premièrement, ils donnent 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.

3.4.2.2. Statistiques locales et cartes de significativité de Moran

La statistique globale de Moran n'est autre que la moyenne des statistiques locales de Moran (Longley et Tobon, 2004). En effet, Anselin (1995) décompose l'indice de Moran en isolant la part de chaque individu dans la mesure globale. L'intérêt de cette démarche est de proposer un lien direct de proportionnalité entre les indices locaux et l'indice global (Anselin, 1995 et 1996 ; Oliveau, 2005 ; Ayadi et Amara, 2009).

Les indices d'autocorrélation spatiale locale représentent des statistiques de "*scan spatial*" dans la mesure où ils produisent un vecteur de résultats qui se prête à la représentation cartographique. Anselin (1995) fournit une version locale de la statistique de Moran (I) et de celle de Geary (C). En pratique, c'est l'indice de Moran localisé qui est utilisé.

22. A noter que "élevé" (respectivement "faible") signifie au-dessus (respectivement en-dessous) de la moyenne française

Encadré 3.2 : Statistiques locales de Moran

Pour chaque région i l'indice de Moran I s'écrit de la façon suivante :

$$I = \frac{(x_i - \bar{x})}{m_0} \sum_j W_{ij}(x_j - \bar{x})$$

$$\text{Avec } m_0 = \frac{\sum_i (x_i - \bar{x})^2}{N}$$

Où : x_i est l'observation représentant la région i , \bar{x} est la moyenne des observations. La somme des statistiques locales peut s'écrire de la façon suivante :

$$I = \frac{1}{m_0} \sum_i W_{ij}(x_i - \bar{x}) = \frac{1}{m_0} \sum_i \sum_j W_{ij}(x_i - \bar{x})(x_j - \bar{x})$$

Compte tenu de cette définition, la statistique globale de Moran est proportionnelle à la somme des statistiques locales de Moran :

$$I = \sum_i I_i = S_0$$

Pour une matrice de poids standardisée en ligne $S_0 = N$, la statistique globale de Moran devient :

$$I = \left(\frac{1}{N}\right) \sum_i I_i$$

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 voisins.

La statistique de Moran 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)(Encadré 3.3).

Encadré 3.3 : Statistique locale de Getis et Ord (1992)

Il s'agit d'une statistique utilisée par Getis et Ord (1992) pour étudier la significativité des regroupements spatiaux autour de localisations individuelles. Pour ces différentes statistiques, l'inférence est basée sur une approximation normale. Ainsi, cette statistique est compliquée puisque 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). Nous distinguons l'indice G_i qui exclut l'observation i du calcul, de l'indice G_i^* qui l'inclut. L'indice de G_i^* est le plus récemment utilisé. Soit W_δ , la matrice de proximité et Z le caractère d'intérêt, les coefficients d'influence spatiale W_{ij} sont binaires :

$$\epsilon_{NN} = \frac{NN - \mu_{NN}}{\sqrt{V(NN)}}$$

$\forall i \text{ et } j \in N$

$\forall \text{ poids } W_{ij} = 1 \text{ si } d_{ij} \leq \delta$

Dans le cas de G_i , i et j doivent être différents, $W_{ij} = 0$

G_i^* , i et j sont quelconques, donc $W_{ij}^* = 0$

La formule du coefficient local est donc la suivante :

$$G_i(\delta) = \frac{\sum_{j=1}^m W_{ij}(\delta z_j)}{\sum_{j=1}^n z_j}$$

Sous hypothèse nulle, l'espérance de l'indice G_i où W_i désigne le nombre de voisins qui entrent dans le calcul de l'indice et diffère légèrement de l'espérance de l'indice G_i^* .

$$V(G_i(\delta)) = \frac{W_i(n-1-W_i)}{(n-1)^2(n-2)} \times \frac{S_{t2}}{S_{t1}^2}$$

Avec $S_{t1} = \sum_j z_j$ et $S_{t2} = \frac{\sum_j z_j^2}{n-1} - S_{t1}^2$

Comme les indices G sont sensibles à un effet de taille, nous les remplaçons par leur valeur standardisée :

$$\xi(G_i) = \frac{G_i - E(G_i)}{\sqrt{G_i(G_i)}}$$

Cette statistique est d'espérance nulle et de variance unitaire. Nous rejetons H_0 lorsque la valeur absolue de l'indice standardisé ξ dépasse le quantile correspondant de la loi normale centrée réduite $|\xi(G_i)| \geq 1.96$. Une valeur positive de $\xi(G_i)$ indique la localisation des agrégats de valeurs élevées du caractère d'intérêt, et une valeur négative de $\xi(G_i)$ indique la localisation des agrégats de valeurs faibles du caractère.

Ces outils constituent donc un ensemble de techniques visant à décrire et représenter les distributions spatiales, à identifier des localisations atypiques, à examiner des schémas de localisation des associations spatiales, des clusters, ainsi qu'à suggérer des formes d'hétérogénéité spatiale sur une zone d'étude (Haining, 1990 ; Bailey et Gatrell, 1995 ; Anselin, 1998a,b ; Amara, 2010).

L'indice de Moran et celui de Getis sont intéressants pour identifier les singularités locales. Comme l'indice de Moran, l'indice G de Getis et Ord croît en proportion de la dépendance spatiale, mais il croît aussi selon la taille de l'échantillon et l'unité de mesure. C'est pourquoi il est remplacé par sa valeur centrée-réduite, et s'interprète en fonction des quantiles de la loi normale de Laplace-Gauss. Contrairement à l'indice de Moran, l'indice de Getis ne peut être calculé que pour des valeurs positives, typiquement des densités ou des fréquences. En effet, l'indice de Moran est associé à un corrélogramme. Il sert à la fois à des tests d'hypothèse, à la recherche

d'agrégats et à celle de singularités²³. En outre, Boots et Tiefelsdorf (2000) montrent à travers des simulations effectuées sur des configurations spatiales régulières que la distribution pour les statistiques de Moran I ne peut être approximée par une distribution normale. L'inférence statistique doit alors être basée sur l'approche des permutations (Anselin, 1995). Dans ce cas, les probabilités critiques obtenues pour les statistiques locales de Moran sont, en fait, des pseudo-niveaux de significativité. Anselin (1995) donne deux interprétations pour les indicateurs LISA. Premièrement, ils peuvent être utilisés comme indicateurs de regroupements spatiaux locaux, de la même façon que les statistiques de $G_i(d)$. Deuxièmement, ils permettent de diagonaliser l'instabilité locale (les localisations atypiques ou les poches de non-stationnarité), les observations extrêmes et les régimes spatiaux. Cette seconde interprétation est similaire à l'utilisation du diagramme de Moran. En effet, les observations extrêmes pour les LISA seront associées aux régions qui contribuent le plus à l'autocorrélation spatiale globale (Virol, 2006; Amara, 2010). En combinant l'information donnée par le diagramme de Moran et la significativité des LISA, nous obtiendrons des cartes de significativité de Moran. Ces cartes présentent les régions aux associés avec un LISA significatif et indiquent par un code de couleur les quadrants du diagramme de Moran auxquels ces régions appartiennent (Anselin et Bao, 1997). Ceci nous permet de déterminer le degré de ressemblance de chaque département avec les départements voisins.

3.4.3. Une application des outils de mesure de l'autocorrélation spatiale locale sur données d'innovation

Comme nous l'avons évoqué dans la sous section précédente, une des limites du diagramme de Moran est qu'il ne fournit pas d'information sur la significativité des regroupements spatiaux. Celle-ci est obtenue à l'aide d'indicateurs locaux d'association spatiale. Pour cela, nous utilisons dans ce qui suit ces deux outils simultanément afin d'étudier la nature de l'autocorrélation spatiale de l'innovation entre départements français.

3.4.3.1. Détection de l'autocorrélation spatiale locale au niveau des brevets

Pour visualiser les différents types d'association spatiale entre départements, nous construisons les diagrammes de Moran associés aux brevets pour les années 2002 et 2008 :

23. L'indice de Geary lui ressemble, mais il identifie plus les régions homogènes que des agrégats.

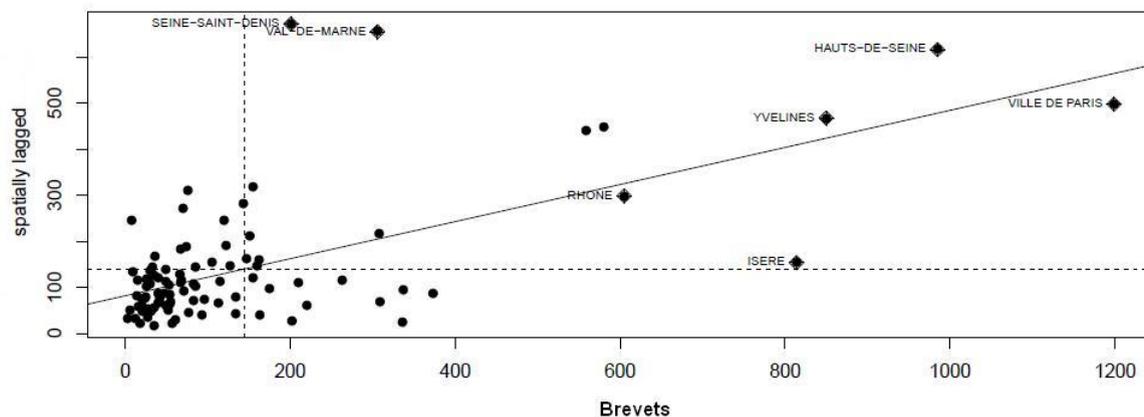


Figure 3.8 – Diagramme de Moran des brevets en 2002

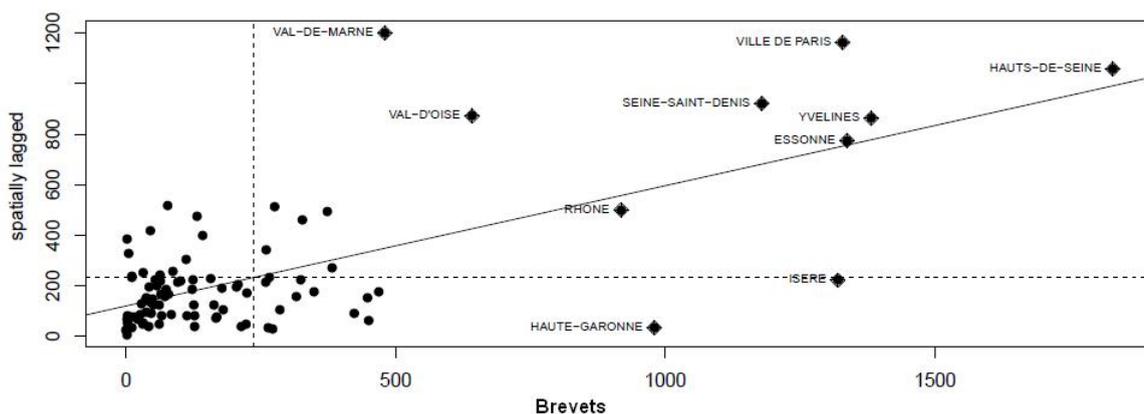


Figure 3.9 – Diagramme de Moran des brevets en 2008

Ces deux diagrammes construits à partir de la matrice de contiguïté d'ordre 1 pour les brevets en 2002 et 2008²⁴, illustrent bien les quatre types d'associations spatiales (H-H ; L-L ; H-L et L-H).

Nous pouvons souligner des discontinuités spatiales et des associations spatiales aussi bien positives que négatives des 94 départements français. Il apparaît que la plupart des départements français sont caractérisés par une association spatiale positive (H-H et L-L). Ainsi, en 2002, 68,08% des associations (75,53% en 2008) appartiennent aux quadrants H-H et L-L. Les diagrammes de Moran permettent également de détecter les régions atypiques (quadrant L-H et H-L). En 2002, 31,92% des départements (24,47% en 2008) dévient du schéma global d'association spatiale (11,7% dans le quadrant H-L et 15,95% dans le quadrant L-H).

24. Le diagramme de Moran peut être cartographié en construisant une carte appelée "la carte de Moran" qui permet de présenter toutes les associations spatiales significatives et non significatives.

Néanmoins, ces associations ne peuvent pas être interprétées puisque nous n'avons aucune information sur leurs significativités. Un test de permutation, basé sur un tirage aléatoire de type Monte Carlo, est proposé par Anselin (1995). Les p-values inférieures ou égales au seuil de 5% ont ainsi été retenues pour réaliser les cartes suivantes :

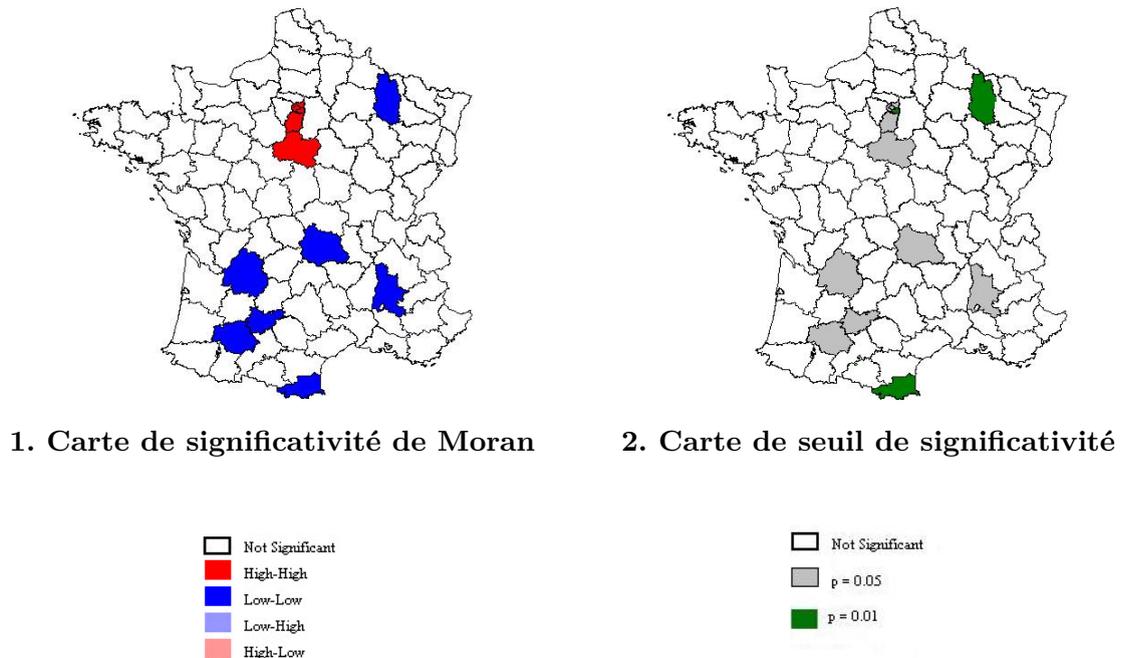


Figure 3.10 – Cartes de significativité de Moran des brevets en 2002

L'analyse de l'autocorrélation spatiale au niveau local par le biais du diagramme de Moran ne peut être interprétée correctement qu'avec les cartes de significativité de Moran. Cette significativité (à 1% et à 5%) atteste qu'une valeur associée à un département i est significativement influencée par la valeur des départements voisins, définis par la matrice de poids spatiale. Les cartes de significativité de Moran ci-dessus mettent en évidence des *clusters* d'innovation pour l'année 2002 (carte n°1 de la figure 10) ainsi que leurs valeurs significatives associées (carte n°2 de la figure 10).

En effet, nous observons sur la figure 10 un *cluster* positif (H-H), à valeur supérieure à la moyenne et hautement significative et composé des départements Val-de-Marne (94), Hauts-de-Seine (92), Loiret (45), Seine-Saint-Denis (93), Essonne (91) et Paris (75), et 7 *clusters* positifs (L-L), à valeur inférieure à la moyenne et significative au seuil de 5% (carte n°2 de la figure 10). A travers la figure 10, nous remarquons aussi que seulement les quadrants les H-H et les L-L sont significatifs. Ainsi, il n'y a donc pas d'agglomération atypique significative.

La liste des *clusters*, les statistique de LISA ainsi que les p-value correspondantes sont représentées dans le tableau suivant :

Tableau 3.6 – Statistiques de LISA des brevets en 2002

Nom du département	code	Statistiques LISA	<i>Clusters</i>	p-value
Val-De-Marne	94	1.261	H-H	0.006
Meuse	55	0.329	L-L	0.01
Hauts-De-Seine	92	3.914	H-H	0.014
Pyrénées-Orientales	66	0.307	L-L	0.018
Loiret	45	0,07217	H-H	0.02
Seine-Saint-Denis	93	0.277	H-H	0.022
Gers	32	0.299	L-L	0.024
Tarn-Et-Garonne	82	0.258	L-L	0.04
Drome	26	0.168	L-L	0.042
Essonne	91	1.434	H-H	0.046
Dordogne	24	0.261	L-L	0.048
Puy-De-Dome	63	0.110	L-L	0.048
Paris	75	4.896	H-H	0.05

En comparaison avec la carte des *clusters* des brevets de 2008, nous constatons que la structure spatiale départementale a gagné, depuis, des départements qui présentaient des dynamiques positives en Île-de-France et au sud de la France, il apparaît aussi un département où l'autocorrélation spatiale est atypique (H-L), la Haute-Garonne :

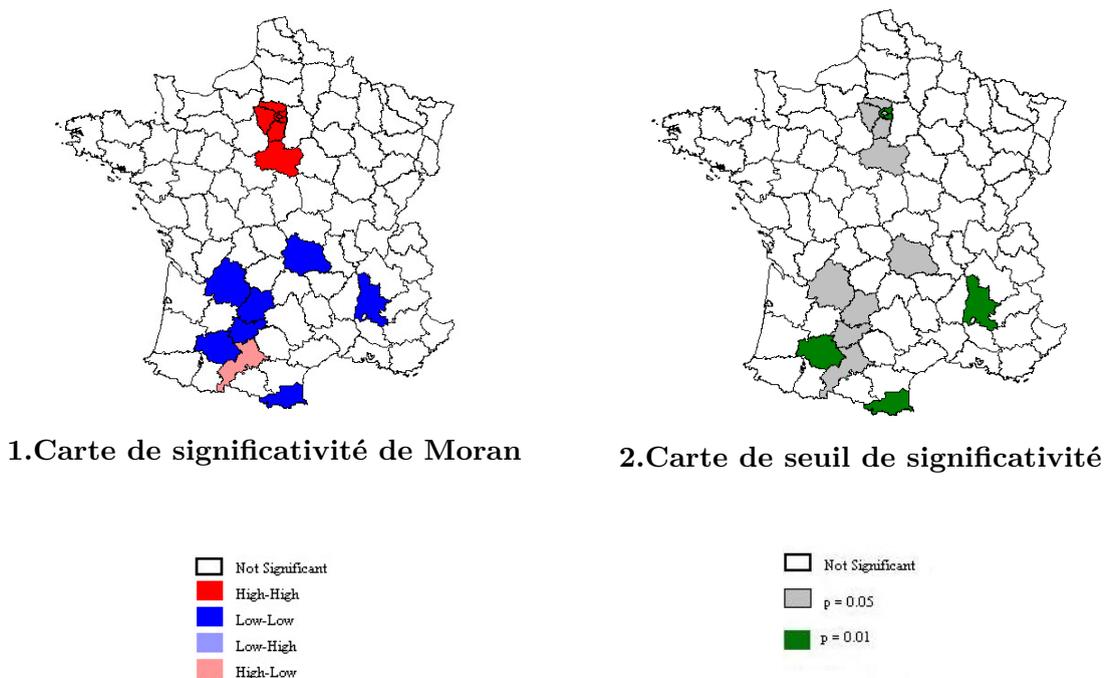


Figure 3.11 – Cartes de significativité de Moran des brevets en 2008

Tableau 3.7 – Statistiques de LISA des brevets en 2008

Nom du département	code	Statistiques LISA	Clusters	p-value
Val-De-Marne	94	1.536	H-H	0.005
Hauts-De-Seine	92	4.328	H-H	0.01
Seine-Saint-Denis	93	3.726	H-H	0.01
Drome	26	0.201	L-L	0.01
Gers	32	0.342	L-L	0.01
Pyrenées-Orientales	66	0.424	L-L	0,01
Yvelines	78	4.497	H-H	0.02
Val-D'oise	95	1.115	H-H	0.02
Paris	75	4.562	H-H	0.025
Essonne	91	2.841	H-H	0.025
Dordogne	24	0.316	L-L	0.025
Puy-De-Dome	63	0.033	L-L	0.03
Lot	46	0.301	L-L	0.035
Loiret	45	0.059	H-H	0.04
Haute-Garonne	31	-0.955	H-L	0.05
Tarn-Et-Garonne	82	0.218	L-L	0.05

Le tableau (3.7) montre que les départements Val-D'oise (95) et Yvelines (78) présentent une autocorrélation positive avec leurs voisins alors que ce n'était pas le cas en 2002. Ces *clusters* d'innovation sont représentés par une valeur *H-H* (en rouge) puisqu'ils présentent un niveau élevé de brevets tout en étant directement entourés par des départements voisins présentant également un niveau élevé de brevets. Dans ce quadrant, nous remarquons que la région Île-de-France est un "*hotspot*" majeur d'activités d'innovation en France. Elle rassemble à elle seule 36% des brevets français en 2002 et 38% en 2008 ce qui laisse penser que la région Île-de-France constitue le coeur de l'innovation en France. A l'échelle européenne, elle constitue la première région innovante. Le département Loiret se situe aussi dans le quadrant *H-H* indiquant une autocorrélation spatiale positive au dessus de la moyenne et entouré par des départements à niveau élevé de brevets notamment le département Essone (91) de la région Île-de-France.

Par ailleurs, dans le quadrant L-L nous trouvons le département de la Meuse (55), le département du Puy-de-Dôme (63), Drôme (26), la Dordogne (24), le Tarn-et-Garonne (82), le Gers (32) et les Pyrénées-Orientales (66).

En 2008, le département Lot (46) se rajoute au quadrant *L-L* (en bleu) pour former un *cluster* présentant une autocorrélation positive de valeurs inférieures à la moyenne. Le nombre de brevets moyen entre 2002 et 2008 varie entre 14 pour le département Puy-de-Dôme et 165 brevets pour le département Essone (91).

Nous précisons à ce niveau que les quadrants de type L-L ne représentent pas nécessairement les départements ayant le nombre de brevets le plus bas. En réalité, c'est le département de la Lozère qui présente le nombre le plus bas de brevets. Néanmoins, ce département ne s'affiche pas sur la carte de significativité de LISA. Ceci est expliqué par le fait que ce département ne présente aucune corrélation spatiale significative avec ses départements voisins. Ainsi, même si le nombre de brevet est faible dans un département donné, ce dernier ne peut pas être significativement autocorrélé spatialement avec un département voisin. C'est ici que nous comprenons la différence entre concentration et agglomération (autocorrélation spatiale).

Le département Haute-Garonne présente une position atypique (H-L), il affiche des valeurs fortes de brevets alors qu'il est entouré par des voisins qui présentent des valeurs relativement faibles. La DATAR lui attribue même la première place en matière de potentiel scientifique et technique dans son rapport de février 2004. Cette région fournit donc un exemple type de ce que Krugman (1996) appelle le phénomène "*d'ombre d'agglomération*" ce qui traduit la domination que peut exercer un centre économique important sur le reste de son territoire plus précisément les voisins.

Nous vérifions dans ce qui suit si les dépenses en R&D présentent une structure d'autocorrélation différente de celle des brevets.

3.4.3.2. Autocorrélation spatiale au niveau des dépenses en R&D

Nous présentons dans la figure 12 les cartes de significativité de Moran des dépenses en R&D interne pour les années 2002 et 2008. Les cartes de significativité de Moran pour les dépenses en R&D externe publique et privée sont présentées dans la carte (3.12) :

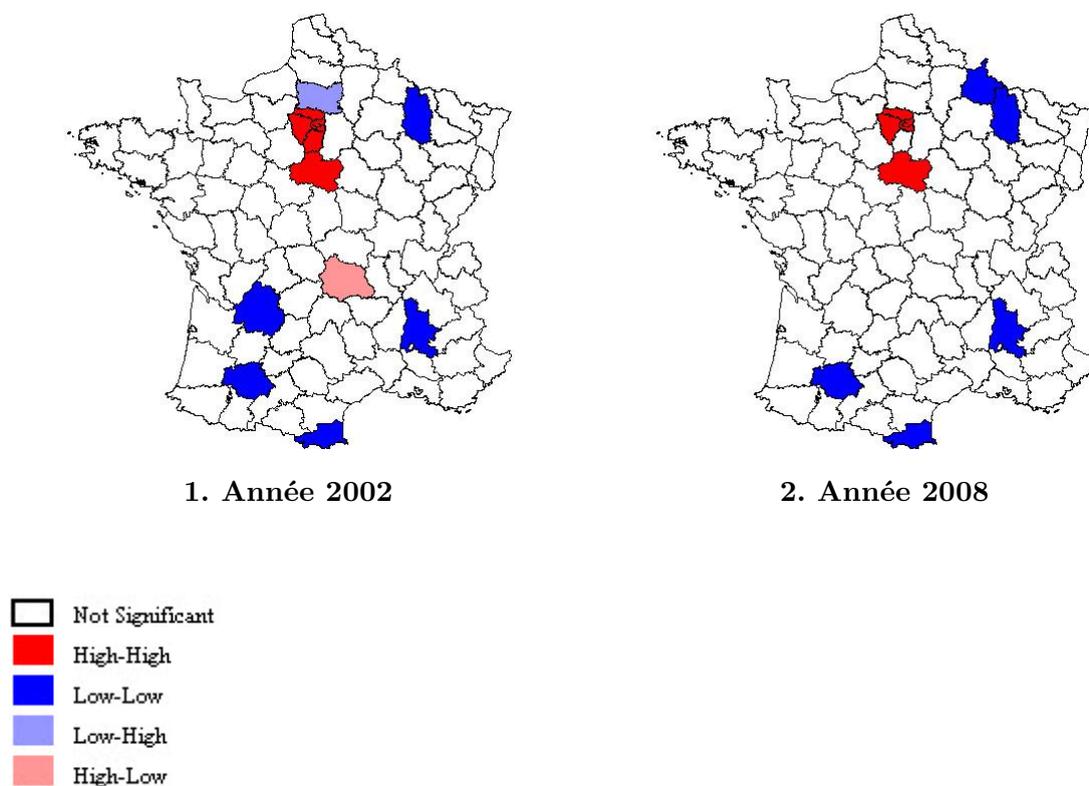


Figure 3.12 – Cartes de significativité de Moran des dépenses en R&D interne en 2002 et 2008

Les cartes de significativité de Moran révèlent un haut niveau d'hétérogénéité dans les schémas de localisation des dépenses en R&D interne.²⁵

En 2002, le quadrant *H-H* est représenté par les départements de la région Île-de-France (sauf le département Seine-et-Marne) et le département Loiret. Ces départements révèlent des dépenses locales fortes en R&D interne tout en étant entourés par des départements à fort niveau de dépenses en R&D interne.

Le quadrant *L-L* regroupe 5 départements : la Meuse (55), la Drome (26), le Dordogne (24), le Gers (32) et les Pyrénées-Orientales (66).

L'étude de l'autocorrélation spatiale des dépenses en R&D interne révèle une structure d'autocorrélation spatiale proche de celles des brevets. En effet, ces mêmes départements représentent aussi une autocorrélation spatiale positive et au-dessus de la moyenne pour le cas des brevets. Ceci peut être expliqué par la relation étroite qui existe entre les dépenses en R&D interne et les brevets. En effet, les dépenses en R&D interne sont considérées comme un facteur primordial dans le processus d'innovation.

25. Les statistiques de LISA sont présentées dans l'annexe 3.

Cette relation de causalité est aussi manifestée à travers la relation qui existe entre le lieu de l'investissement en R&D et le lieu de production de l'innovation. Ceci est moins vrai pour le quadrant L-L de l'année 2008. En effet, la relation de causalité est moins vraie pour le cas des départements témoignant d'une dynamique très faible d'investissement en R&D et de production d'innovation. Nous ne pouvons pas conclure à ce niveau qu'une autocorrélation spatiale liée à une dynamique très faible en R&D est une source d'autocorrélation spatiale liée à une dynamique faible de brevets. En effet, la relation R&D-innovation n'est pas toujours réciproque dans le cas d'une analyse d'autocorrélation spatiale.

En plus des quadrants H-H et L-L, les quadrants H-L (en rose) et L-H (en bleu clair) sont aussi représentés pour le cas de la carte de significativité de Moran des dépenses en R&D interne. Le département du Puy-de-Dôme (63) est représenté par le quadrant H-L. C'est un département à dépenses élevées en R&D interne mais qui est entouré par des voisins à dépenses en R&D faibles. Par contre, le département d'Oise est situé dans le quadrant L-H. Même si les dépenses en R&D interne du département Oise (213.602 millions d'euros) sont importantes par rapport à la moyenne (235.180 millions d'euros), ce dernier est considéré comme un département à valeur faible par rapport à ses voisins les plus proches et autocorrélés négativement avec ses voisins de la région Île-de-France (ayant des dépenses les plus élevées en R&D interne en France). Ces configurations peuvent s'interpréter comme des *spillovers* négatifs associés à une concurrence entre départements français. En 2008, nous remarquons une légère baisse du phénomène de dépendance spatiale entre départements. D'ailleurs, les quadrants L-H et H-L disparaissent et le nombre de départements H-H a diminué.

3.4.3.3. Autocorrélation spatiale au niveau des dépenses en R&D externe

L'autocorrélation spatiale locale positive et significative entre départements devient moins remarquable avec les dépenses en R&D externe. Contrairement à la R&D interne, la R&D externe contient une proportion plus élevée de connaissances codifiées que de celles tacites. Ainsi, leur effet spatial est moins prononcé entre départements puisque les connaissances qui découlent de la R&D externe ne nécessitent pas une proximité spatiale pour se diffuser, elles peuvent se transmettre via les réseaux d'innovation utilisant les TIC par exemple. Nous pensons que cela provient du fait que la R&D externe contient une faible proportion de connaissances tacites et se base beaucoup plus sur des connaissances codifiées moins génératrices d'externalités.

Nous avons distingué les cartes de significativité de Moran de la R&D externe publique de celle privée afin de mieux illustrer l'hétérogénéité spatiale entre départements :

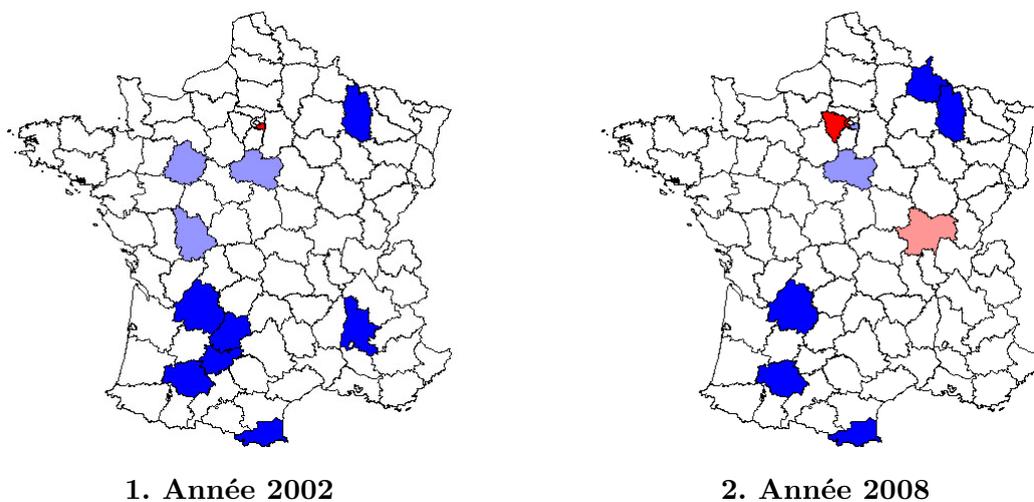


Figure 3.13 – Cartes de significativité de Moran des dépenses en R&D externe publique en 2002 et 2008

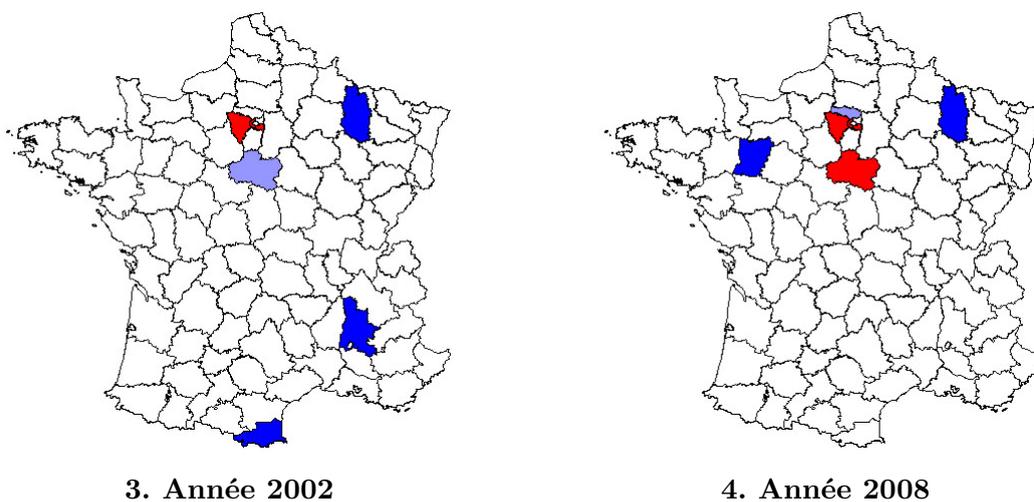


Figure 3.14 – Cartes de significativité de Moran des dépenses en R&D externe privée en 2002 et 2008



Nous remarquons que les dynamiques spatiales des dépenses en R&D externes publiques en 2002 sont non significatives pour le cas de la région Île-de-France (sauf

pour le département du Val-de-Marne). Nous identifions aussi des départements à comportement spatial atypique représentant le quadrant L-H : Loiret (45) et Vienne (86). Ces départements se distinguent par des valeurs faibles de dépenses en R&D publique en comparaison avec leurs voisins²⁶.

En 2008, le nombre de départements dans le quadrant L-H baisse et d'autres départements apparaissent significatifs et à autocorrélation spatiale positive (Yvelines (78), Saône-et-Loire (71) et Hauts-de-Seine (92)). Ceci peut être expliqué par l'augmentation des dépenses en R&D publiques en comparaison avec l'année 2002. Les externalités de connaissance qui découlent des départements influent positivement sur les voisins. Par ailleurs, les cartes de significativité de Moran des dépenses en R&D externe privée montrent une faible autocorrélation spatiale entre les départements. En effet, le nombre de départements significatifs en 2002 est très peu élevé et il n'existe que 3 départements dans le quadrant L-L, 1 département dans le quadrant L-H et 2 dans le quadrant H-H. En 2008, nous n'observons que 2 départements dans le quadrant L-L, 3 dans le quadrant H-H et 1 dans le quadrant L-H. L'autocorrélation spatiale est donc très faible entre les départements pour le cas des dépenses en R&D externes. Globalement, les cartes de significativités de Moran des brevets et des dépenses en R&D présentent un nombre relativement faible de départements autocorrélés spatialement avec leurs voisins (moins de 1/4 des 94 départements métropolitains). Ceci montre que l'effet spatial, même s'il existe, est faible. Le transfert des connaissances entre départements ne se limite pas seulement à l'espace et aux externalités localisées de connaissance. D'autres facteurs non spatiaux sont alors plus importants dans la transmission des connaissances et par la même occasion dans la production de l'innovation.

3.5. Conclusion

L'objectif de ce chapitre était de caractériser les dynamiques spatiales observées pour l'innovation en 2002 et 2008. Dans un premier temps, nous avons mesuré la concentration des activités innovantes en utilisant l'indice de Gini. Les résultats montrent que les brevets et les dépenses en R&D interne et externe sont concentrés (Indice de Gini entre 0.6 et 0.84), mais cette approche ne nous dit pas dans quels départements s'effectue cette concentration. En outre, la position des unités spatiales ainsi que la distance qui les sépare ne sont pas prises en compte dans le calcul de l'indice de concentration de Gini, ni dans les autres indices de concentration (Herfindhal, Elison Glaeser, etc). C'est pourquoi, nous nous focalisons dans un deuxième temps sur l'étude

26. Voir l'annexe 3 pour plus de détails sur les significativités des statistiques de LISA.

de l'agglomération des activités innovantes afin de mettre en lumière les dynamiques distinctes selon les départements.

Nous avons mobilisé les outils de l'AEDS (analyse exploratoire des données spatiales) afin de tester la significativité statistique des concentrations et d'identifier leurs schémas de localisation. Ces outils sont regroupés dans deux groupes. Le premier groupe permet de mesurer l'agglomération globale à travers l'utilisation de la statistique de Moran globale et le deuxième s'intéresse à l'analyse locale de l'autocorrélation entre départements. Les outils mobilisés pour ce dernier groupe sont le diagramme de Moran, les statistiques de LISA puis les cartes de significativité de LISA, qui est une combinaison des deux derniers indicateurs locaux. L'analyse par le diagramme de Moran nous permet de repérer les différents régimes spatiaux et de déterminer les localisations qui correspondent ou non au schéma global d'autocorrélation spatiale (Guillain et al., 2004; Amara, 2010). Les statistiques de LISA nous ont permis de mesurer les autocorrélations spatiales locales pour chaque département ainsi que leurs niveaux de significativité. L'utilisation conjointe des statistiques de LISA et du diagramme de Moran nous ont permis d'identifier les clusters innovants à autocorrélation spatiale significative. Le calcul de ces indicateurs (au niveau global et local) a nécessité le choix d'une matrice de pondération. Ce choix s'est révélé crucial puisqu'il a conditionné très fortement les résultats obtenus et par là même l'analyse qui en a découlé.

Nos résultats montrent que les dynamiques positives et significatives observées pour le cas des brevets et des dépenses en R&D sont surtout dans la région Île-de-France. En effet, l'ensemble de l'analyse laisse apparaître la persistance de disparités relativement importantes au sein des départements français qui se matérialisent par des concentrations différentes pour les années 2002 et 2008. Il existe une logique sous-jacente à la dynamique d'agglomération qui n'est pas complètement aléatoire, ceci mettrait en évidence des externalités de connaissance qui sont, en partie, le résultat des agglomérations. Ces externalités positives de connaissance sont plus importantes pour le cas de la région Île-de-France. En effet, l'ensemble de l'analyse montre que cette région est au centre de dynamiques positives, en terme de production d'innovation puisqu'elle occupe la première place au niveau du nombre de brevets (Voir figure 2.7 du chapitre 2 pour une répartition des brevets par département entre 2002 et 2008).

Néanmoins, le niveau départemental trop agrégé pourrait être une explication de la forte autocorrélation spatiale positive qui est manifeste au niveau des départements de la région Île-de-France²⁷.

27. Une extension de ce travail, uniquement centré sur la région Île-de-France, sera prochainement menée pour qualifier la structure spatiale de la région au regard de la production d'innovation.

Les effets de débordement et les externalités spatiales à l'origine de l'autocorrélation spatiale sont donc susceptibles d'influencer les processus d'innovation départementaux et devraient par conséquent être sérieusement prises en compte. Par ailleurs, les résultats jusqu'à présent ne nous permettent pas de nous prononcer sur l'ampleur de l'effet spatial sur la production d'innovation. Ils nous ont seulement permis de détecter la présence de plusieurs *clusters* d'innovation autocorrélés spatialement qui demandent confirmation. La poursuite de notre travail consistera à mesurer l'effet spatial sur la production de l'innovation à travers une modélisation économétrique spatiale.

Chapitre 4

De l'intérêt d'une analyse économétrique spatiale appliquée à des données de panel

4.1. Introduction

Ce chapitre a pour objectif de mesurer de manière plus directe le rôle de la proximité géographique dans le processus d'innovation au niveau des 94 départements français métropolitains (Corse exclue). Autrement dit, nous proposons dans ce dernier chapitre une étude complémentaire à celle réalisée dans le chapitre 3, à travers l'utilisation des modèles économétriques spatiaux.

L'introduction de l'espace dans les modèles économétriques vise à étudier la présence de deux effets spatiaux importants : l'autocorrélation spatiale qui se réfère à l'absence d'indépendance entre observations géolocalisées, et l'hétérogénéité spatiale¹ qui est liée à la différenciation des variables et des comportements dans l'espace du fait que 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, etc. (Anselin, 1988 ; Le Gallo, 2002a ; Jayet, 2001).

Jusqu'à une période relativement récente, les études empiriques tentant de capturer les éléments géographiques et spatiaux sont assez rares. Cela est dû essentiellement à la non disponibilité des données à des échelles spatiales plus fines, d'une part, et à la difficulté de mesurer explicitement ces externalités spatiales, d'autre part. Par ailleurs, ignorer les interactions spatiales dues à la proximité géographique, si elles existent, provoque des erreurs dans les estimations et mène notamment à des paramètres

1. Voir chapitre 3 pour plus de détails sur le concept d'hétérogénéité spatiale.

d'intérêt biaisés. Il est donc important de prendre en compte les interactions spatiales dans les modèles économétriques modélisant les données d'innovation.

Pour tester la présence de la dépendance spatiale, trois modèles sont proposés. Autrement dit, les effets spatiaux peuvent être incorporés dans un modèle de régression linéaire de plusieurs manières (Blommestein, 1983; Anselin, 1988 : 35-39; Florax et Folmer, 1992). Une première manière consiste à spécifier un modèle spatial autoregressif (spatial autoregressive model, SAR), intégrant la variable endogène décalée spatialement. Une deuxième manière consiste à intégrer un processus autorégressif sur les erreurs (spatial error model, SEM). Il s'agit d'une structure d'autocorrélation sur les erreurs dans la mesure où c'est le terme d'erreur qui englobe la dépendance spatiale (Baltagi et Li, 2006). Enfin, une troisième manière consiste à tenir compte explicitement de l'effet spatial : il s'agit d'incorporer dans le modèle, à la fois, des variables explicatives et une variable à expliquer décalées spatialement (spatial durbin model, SDM).

Nous pouvons aussi considérer une spécification mixte incorporant les deux formes de dépendance spatiale (Florax et Folmer, 1992; Kelejian et Prucha, (1998, 1999)). Nous trouvons par exemple des processus beaucoup plus généraux comportant plusieurs décalages et/ou comportant à la fois une variable autorégressive et une autocorrélation des erreurs (modèle SARAR et modèle SARMA) (Brandsma et Kelletaper, 1979; Huang, 1984; Anselin, 1988 et Jayet, 1993).

Toutefois, peu d'études empiriques dans la littérature des modèles spatiaux ont proposé des modèles incluant à la fois un processus autorégressif et un modèle à autocorrélation des erreurs. Selon Le Gallo (2002a) : "*on retient rarement un modèle avec les deux types d'effets spatiaux, on cherche à modéliser la dépendance spatiale, soit par l'autocorrélation des erreurs, soit par une variable spatiale autorégressive, et non les deux*" (P.15). Le choix se porte donc essentiellement sur le modèle autorégressif ou le modèle avec autocorrélation des erreurs.

Dans cette perspective, nous proposons dans ce chapitre deux modèles autorégressifs (SAR et SDM) et un modèle à autocorrélation spatiale au niveau des erreurs (SEM). L'estimation de ces modèles spatiaux repose sur la notion de contiguïté. Cette contiguïté est exprimée formellement dans une matrice afin de définir les externalités spatiales des activités de recherche, causées par les différents types de proximité entre régions (Anselin, 1988). Nous avons déjà décrit ces matrices au chapitre précédent.

Par ailleurs, il sera question par la suite de l'interprétation des paramètres des modèles contenant des décalages spatiaux de la variable dépendante. En effet, plusieurs auteurs ont montré que ces modèles spatiaux nécessitent une interprétation particulière

des coefficients des variables explicatives (Anselin, 1988 ; Kelejian et al., 2006 ; Kim et al., 2003 ; Le Gallo et al., 2003 ; Amara et al., 2010). De même qu'un décalage temporel nécessite la distinction entre les effets immédiats et les effets cumulés dans le temps, un décalage spatial nécessite de calculer l'effet cumulé de la variable en prenant en compte des effets réciproques ou des "*feedbacks loops*" (voir section 4.5) des départements voisins. Ainsi, une correction proposée par LeSage et Pace (2009), consiste à décomposer les effets spatiaux en effets directs (effets sur la région elle-même y compris les effets réciproques) et en effets indirects (effets sur les autres régions).

Le chapitre s'organise de la manière suivante. Dans la section 2, nous présenterons les modèles spatiaux utilisés pour estimer nos données d'innovation. Dans la section 3, nous discuterons la méthode d'estimation. La section 4 sera consacrée à la présentation des résultats d'estimation des modèles spatiaux. La section 5 présentera les résultats de la décomposition de l'effet spatial en effets directs et effets indirects. Enfin, la section 6 conclut le chapitre.

4.2. Modèles économétriques spatiaux sur données de panel

Nous consacrons cette section à la présentation des modèles les plus célèbres d'économétrie spatiale². Ces modèles de régression spatiale représentent des cas particuliers de modèles d'équations simultanées destinées à décrire les répartitions d'un indicateur y en un lieu, en fonction des variables explicatives et de ses variations dans les autres lieux du domaine (d'Aubigny, 2006).

La prise en compte de l'autocorrélation spatiale dans les modèles économétriques peut s'effectuer de plusieurs manières : par des variables spatiales décalées (endogènes ou exogènes) ou par une autocorrélation spatiale des erreurs (Le Gallo, 2002b). Comme point de départ, nous retenons le modèle de régression linéaire classique suivant :

$$y_{it} = x'_{it}\beta + \mu_i + \sum_t time_t + \epsilon_{it}, \quad i = 1, \dots, N \quad et \quad t = 1, \dots, T \quad (4.1)$$

Nous retenons les conventions d'écriture suivantes :

N est le nombre de départements et T la durée de l'étude. y_{it} est la valeur de la variable dépendante du département i à l'instant t . x_{it} est un vecteur $k \times 1$ d'observations relatives aux k variables explicatives et β est un vecteur de paramètres de dimension $k \times 1$. μ_i est l'effet individuel, spécifique au département i , $time_t$ représente l'effet temporel à l'instant t . Les ϵ_{it} sont supposées indépendantes des effets μ_i et des variables

2. Nous nous limitons au cas des répartitions gaussiennes des résidus.

explicatives du modèle. Elles sont supposées satisfaire les deux hypothèses suivantes : H_1 : les erreurs ϵ_{it} sont identiquement et indépendamment distribuées (*iid*), suivant une loi normale, d'espérance nulle et de variance finie σ_ϵ^2 ($\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$).

H_2 : la matrice X est non-stochastique, de rang complet k et $\lim_{n \rightarrow \infty} (1/N)X'X = Q$ où Q est une matrice finie et non-singulière.

Partant de ce modèle a-spatial, nous construisons trois spécifications afin de prendre en compte l'autocorrélation spatiale entre les observations. Il est utile de préciser que si l'autocorrélation spatiale n'est pas prise en compte, elle peut à la fois générer une mauvaise estimation des coefficients et de leur significativité, mais aussi des erreurs d'interprétation dans les diagnostics standards de régression (Anselin, 1988).

4.2.1. Modèle Spatial Autorégressif (SAR)

Une première façon d'incorporer l'autocorrélation spatiale dans un modèle est la suivante :

$$y_{it} = \rho \sum_{j=1}^N w_{ij} y_{jt} + x_{it} \beta + \mu_i + \sum_t time_t + \epsilon_{it}, \quad (4.2)$$

où y_{it} est la valeur de la variable dépendante ($\ln(bvt_{it})$) du département i à l'instant t . x_{it} est le vecteur ligne $1 \times k$ des variables indépendantes ($\ln(R\&Dinterne)$, $\ln(R\&Dextpub)$, $\ln(R\&Dextpiv)$, $\ln(EGindex)$)' $_{it}$. w_{ij} est l'élément (i, j) de la matrice de pondération spatiale W , définie telle que $\sum_{j=1}^N w_{ij} = 1$ afin que W soit standardisée. ρ est le coefficient d'autorégression spatiale indiquant l'intensité de l'interaction existante entre l'observation y_{it} et ses voisins. β est de taille $k \times 1$. μ_i représente l'effet individuel du département i , $time_t$ représente l'effet temporel à l'instant t et ϵ_{it} est le terme d'erreur.

Dans ce modèle, l'observation y_{it} est en partie expliquée par la moyenne des valeurs prises par la variable dépendante au voisinage du département i ³. Cette quantité est représentée par le terme $\sum_{j=1}^N w_{ij} y_{jt}$, qui correspond à la variable dépendante décalée spatialement (*Spatial Lag*) (Encadré 4.1). L'introduction de la variable dépendante spatialement décalée dans (4.1) permet donc d'apprécier le degré de dépendance spatiale et les effets de débordements lorsque les autres variables sont contrôlées. Ainsi, cette variable permet d'évaluer l'impact des autres variables explicatives tout en contrôlant la dépendance spatiale entre les observations. Dans la littérature, le coefficient de la variable Wy mesure l'effet de voisinage et

3. Nous parlons ici de moyenne car la matrice W est standardisée. Cette standardisation facilite aussi la comparaison de l'ampleur de ρ lorsque (4.2) est estimée pour plusieurs matrices de poids.

l'intensité de l'interaction entre les différents départements dans l'échantillon considéré.

Encadré 4.1 : Construction de la variable *Spatial Lag* ou Wy

Afin de mieux comprendre la construction de la variable Wy , nous considérons la matrice de contiguïté d'ordre 1 (C) qui mesure la contiguïté des départements de la région Midi-Pyrénées. Pour simplifier, nous nous plaçons ici dans un exemple de données en coupe.

La région Mid-Pyrénées regroupe huit départements : l'Ariège (09), l'Aveyron (12), la Haute-Garonne(31), le Gers(32), le Lot(46), les Hautes-Pyrénées (65), le Tarn(81) et le Tarn-et-Garonne (82). Si deux départements ont une frontière commune, l'élément de la matrice prend la valeur 1 et 0 sinon (les caractéristiques des matrices de contiguïtés ont été déjà présentées dans le chapitre 3).

$$C = \begin{pmatrix} & \mathbf{09} & \mathbf{12} & \mathbf{31} & \mathbf{32} & \mathbf{46} & \mathbf{65} & \mathbf{81} & \mathbf{82} \\ \mathbf{09} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{12} & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ \mathbf{31} & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ \mathbf{32} & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ \mathbf{46} & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ \mathbf{65} & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ \mathbf{81} & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ \mathbf{82} & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \end{pmatrix}$$

La matrice C est par la suite standardisée. Nous obtenons donc une matrice standardisée notée W . Par la suite, W est multipliée par un vecteur y de dimension $(7, 1)$. Le produit $W \times y = Wy$ est un vecteur de dimension $(7,1)$ représentant le vecteur "*Spatial Lag*" correspondant à chaque observation i , $i = 1, \dots, 7$.

$$Wy = \begin{pmatrix} & \mathbf{09} & \mathbf{12} & \mathbf{31} & \mathbf{32} & \mathbf{46} & \mathbf{65} & \mathbf{81} & \mathbf{82} \\ \mathbf{09} & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{12} & 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 \\ \mathbf{31} & 1/5 & 0 & 0 & 1/5 & 0 & 1/5 & 1/5 & 1/5 \\ \mathbf{32} & 0 & 0 & 1/3 & 0 & 0 & 1/3 & 0 & 1/3 \\ \mathbf{46} & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 1/2 \\ \mathbf{65} & 0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ \mathbf{81} & 0 & 1/3 & 1/3 & 0 & 0 & 0 & 0 & 1/3 \\ \mathbf{82} & 0 & 1/5 & 1/5 & 1/5 & 1/5 & 0 & 1/5 & 0 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{pmatrix} =$$

$$\begin{pmatrix} y_3 \\ (y_5 + y_7 + y_8)/3 \\ (y_1 + y_4 + y_6 + y_7 + y_8)/5 \\ (y_3 + y_6 + y_8)/3 \\ (y_2 + y_8)/2 \\ (y_3 + y_4)/2 \\ (y_2 + y_3 + y_8)/3 \\ (y_2 + y_3 + y_4 + y_5 + y_7)/5 \end{pmatrix}$$

Le coefficient obtenu (Wy) permet de mesurer l'effet de l'innovation des départements voisins d'un département i sur l'innovation du département i .

Finalement, l'ampleur de l'interaction entre le vecteur y et le vecteur Wy est représentée par le coefficient d'autorégression spatiale ρ . Dans le cas où le modèle (SAR) n'est pas significatif, il est possible que le processus d'interaction spatiale s'exprime au travers des erreurs. Autrement dit, nous pouvons faire l'hypothèse que les interactions spatiales se produisent entre des variables inobservées qui agissent sur le terme d'erreur.

4.2.2. Modèle avec autocorrélation des erreurs (SEM)

Nous considérons le modèle SEM suivant (Florax et Folmer, 1992; Baltagi et Li, 2006) :

$$y_{it} = x_{it}\beta + \mu_i + \sum_t time_t + \phi_{it} \quad (4.3)$$

Avec $\phi_{it} = \lambda \sum_{j=1}^N w_{ij}\phi_{jt} + \epsilon_{it}$

λ est un coefficient d'autorégression spatiale des résidus et les erreurs ϵ_{it} sont i.i.d, centrées et de variance σ^2 .

L'autocorrélation spatiale des erreurs s'interprète souvent comme un problème dans la spécification du modèle, telle que l'omission de variables pertinentes dans la mesure où l'effet qui n'est pas capté dans les variables explicatives est imputé à un effet de diffusion spatiale dans les erreurs. Omettre à tort une autocorrélation spatiale des erreurs, peut produire des estimateurs non biaisés mais inefficients⁴ (Le Gallo, 2002b).

Le modèle avec autocorrélation spatiale des erreurs a été utilisé pour améliorer les estimations des modèles de prix hédoniques (Pace et Gilley, 1997; Dubin, 1998; Dubin et al, 1999), ou de convergence conditionnelle (Fingleton, 1999). Il a également permis de modéliser la diffusion d'un choc aléatoire portant sur le PIB d'une région vers les

4. C'est-à-dire un estimateur à variance non minimale.

autres régions aux États-Unis (Rey et Montouri, 1999) ou en Europe (Baumont et al, 2001).

Par ailleurs, d'autres variables peuvent être introduites dans le modèle afin de tenir compte de l'existence des externalités de connaissance locales en considérant les caractéristiques des départements proches. Ainsi, un département peut bénéficier des externalités de connaissance lorsqu'il est situé à proximité des départements ayant une activité de R&D intense qui se manifeste à travers un nombre élevé de demandes de brevets, des dépenses importantes en R&D, de multiples collaborations et alliances en R&D. Ces externalités sont estimées à travers des variables explicatives décalées spatialement et une variable à expliquer décalée spatialement. Le modèle de Durbin spatial (SDM) est donc mobilisé pour cette fin.

4.2.3. Modèle régressif croisé (SDM)

Le modèle SDM contient, outre la variable endogène spatialement décalée, les variables explicatives spatialement décalées (LeSage et Pace, 2009).

$$y_{it} = \rho \sum_{j=1}^N w_{ij} y_{jt} + x_{it} \beta + \sum_{j=1}^N w_{ij} x_{jt} \theta + \mu_i + \sum_t time_t + \epsilon_{it} \quad (4.4)$$

Dans ce modèle, l'observation y_{it} est expliquée par la moyenne des logarithmes des brevets effectués dans les départements voisins à i ($\sum_{j=1}^N w_{ij} y_{jt}$), par la moyenne des variables explicatives des voisins à i ($\sum_{j=1}^N w_{ij} x_{jt} \theta$), par l'effet individuel spécifique du département i (μ_i) et par l'effet temporel $time_t$ à l'instant t .

Le modèle SDM peut se présenter comme une version réduite du modèle SEM et une extension du modèle SAR. En effet, sous l'hypothèse $H_0 : \theta = 0$, le modèle prend la forme d'un modèle SAR. Dans le cas où $\theta = -\rho\beta$, le modèle SDM prend la forme d'un modèle SEM. Si les deux hypothèses ($H_0 : \theta = 0$ et $H_0 : \theta = -\rho\beta$) sont rejetées, le modèle SDM doit être donc retenu (Elhorst, 2010a).

Le choix du modèle spatial le plus adéquat pour estimer nos données peut s'effectuer grâce au test du multiplicateur de Lagrange (Anselin et al., 1996 ; Baltagi et al., 2005 ; Amara, 2009 ; Millo et Piras, 2012) que nous présentons dans la section suivante.

4.3. Méthodes d'estimation et règles de choix des modèles spatiaux

Après avoir présenté les trois modèles les plus utilisés dans la littérature en économétrie spatiale, nous précisons dans cette section les méthodes d'estimation dans

le cas des modèles spatiaux sur données de panel (§ 4.3.1), ainsi que les tests permettant de déterminer quel est le modèle spatial (SAR, SEM ou SDM) le plus adapté (§ 4.3.2).

4.3.1. Méthodes d'estimation des modèles spatiaux sur données de panel

L'application de la méthode des moindres carrés ordinaires (MCO) ou généralisées (MCG)⁵ dans l'estimation des modèles avec dépendance spatiale conduit à des estimateurs biaisés et non convergents (Dubin, 1988 et Florax et Folmer, 1992 ; Amara et al., 2010) pour les raisons suivantes :

- Pour le modèle spatial autorégressif, ces estimateurs ne sont pas convergents car quelque soit la distribution de l'erreur ϵ , celle-ci est toujours corrélée avec Wy . Le décalage spatial d'une observation $(Wy)_{it}$ n'est pas seulement corrélé avec le terme d'erreur en i , mais aussi avec tous les termes d'erreurs de toutes les autres localisations.
- Pour le modèle avec autocorrélation spatiale des erreurs, l'estimateur de β est inefficace alors qu'il est sans biais. Ceci est expliqué par le fait que les erreurs ne sont pas sphériques⁶.

Ainsi, pour aboutir à des estimateurs convergents, asymptotiquement normaux et efficaces dans les modèles spatiaux, il faut utiliser d'autres méthodes d'estimations. Une procédure d'estimation appropriée est la méthode de maximum de vraisemblance (ML)⁷.

5. Dubin, 1988 et Florax et Folmer, 1992 montrent que l'application des moindres carrés généralisés (MCG) dans l'estimation des modèles avec dépendance spatiale conduit à des estimateurs biaisés et non convergents.

Considérons la régression classique suivante : $y = X\beta + \epsilon$, on suppose généralement que $E(\epsilon) = 0$ et que la matrice des variances-covariances des erreurs $E(\epsilon\epsilon^T) = \sigma^2 I_n$ où I_n est la matrice d'identité. La différence de l'estimateur MCG avec l'estimateur MCO est que la matrice de covariance des erreurs est maintenant différente d'un multiple de l'identité : $E(\epsilon\epsilon^T) = \sigma^2 \omega$. La matrice de covariance des erreurs devient une matrice positive connue mais l'estimateur par moindres carrés généralisé s'écrit toujours de la même façon que l'estimateur MCO. Ainsi, les estimateurs MCO et MCG ne sont pas convergents (El Ouardighi, 2002).

6. C'est-à-dire $V(\epsilon) \neq \sigma^2$: les erreurs ne sont pas homoscedastiques et $Cov(\epsilon_i, \epsilon_j) \neq 0$: présence de corrélation entre les erreurs

7. Trois méthodes d'estimation des modèles spatiaux ont été développées dans la littérature : la méthode de maximum de vraisemblance, la méthode bayésienne et la méthode des variables instrumentales ou des moments généralisés. Les deux premières méthodes ont été présentées en détail par LeSage et Pace (2009) dans leurs chapitres 3 et 5 du livre "Introduction to spatial econometrics", et Kelejian et Prucha (1998,1999) pour la 3ème méthode. Cette dernière méthode n'exige pas la normalité du terme d'erreur, cependant elle ignore la condition que les paramètres spatiaux soient dans l'intervalle $(1/r_{min}, 1)$ où r_{min} est la valeur propre négative la plus grande en valeur absolue de la matrice de poids.

4.3.1.1. Estimation par le maximum de vraisemblance

La première étude de l'estimation des modèles spatiaux par le maximum de vraisemblance a été proposée par Ord (1975)⁸. Les conditions pour la convergence, l'efficacité et la normalité asymptotique des estimateurs du MV peuvent être dérivées du travail de Heijmans et Magnus (1986a, 1986b) pour le modèle spatial autoregressif et de celui de Magnus (1978) pour le modèle à erreurs autoregressives.

Fonction de vraisemblance pour le cas d'un modèle SAR

La fonction de vraisemblance à maximiser de l'équation (4.2), qui tient compte de l'hétérogénéité de la variable endogène spatialement décalée $\sum_j w_{ij}y_{jt}$, et de l'effet fixe spatial est présentée comme suit :

$$\text{Log}L = -\frac{NT}{2}\log(2\pi\sigma^2) + T\log|I_N - \rho W| - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T (y_{it} - \rho \sum_{j=1}^N w_{ij}y_{jt} - x_i\beta - \mu_i)^2 \quad (4.5)$$

avec I_n la matrice d'identité d'ordre N .

Afin de maximiser cette fonction, il est nécessaire de calculer le déterminant de la matrice $(I_N - \rho W)$ à chaque itération. Ceci complique les calculs numériques lors de l'estimation. Ord (1975) et Pace et Barry (1997) ont proposé donc de calculer ce déterminant en se basant sur les valeurs propres de la matrice W , déterminées séparément avant les calculs numériques. Le déterminant de la matrice $(I_N - \rho W)$ est donné par :

$$\det(I_N - \rho W) = \prod_{i=1}^N (1 - \rho\omega_i), \quad (4.6)$$

où ω_i désigne la $i^{\text{ème}}$ valeur propre de la matrice W . Deux cas se présentent à cette étape : le cas de l'effet spatial fixe et le cas de l'effet spatial aléatoire.

Pour l'estimation de *l'effet fixe spatial* μ_i en fonction des autres paramètres ρ et β , il suffit que la dérivée de $\text{Log}L$ par rapport à μ_i soit égale à 0 (Elhorst, 2010b).

$$\mu_i = \frac{1}{T} \sum_{t=1}^T (y_{it} - \rho \sum_{j=1}^N w_{ij}y_{jt} - x_{it}\beta) \quad (4.7)$$

Pour l'estimation des autres paramètres du modèle (ρ , β et σ), la solution de μ_i est intégrée dans la fonction de vraisemblance (équation 4.5), ce qui donne la fonction de vraisemblance centrée réduite suivante :

8. Nous présentons dans l'annexe 5 le problème principal du jacobien lié à la maximisation de la fonction log-vraisemblance

$$\text{Log}L = -\frac{NT}{2}\log(2\Pi\sigma^2) + T\log|I_N - \rho W| - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T (y_{it}^* - \rho[\sum_{j=1}^N w_{ij}y_{jt}]^* - x_{it}^*\beta)^2 \quad (4.8)$$

où $y_{it}^* = y_{it} - \frac{1}{T} \sum_{i=1}^N y_{it}$ et où $x_{it}^* = x_{it} - \frac{1}{T} \sum_{i=1}^N x_{it}$

Pour le cas de l'effet spatial aléatoire, la fonction de vraisemblance devient :

$$\text{Log}L = -\frac{NT}{2}\log(2\Pi\sigma^2) + T\log|I_N - \rho W| - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T (y'_{it} - \rho[\sum_{j=1}^N w_{ij}y_{jt}]' - x'_{it}\beta)^2 \quad (4.9)$$

où $y'_{it} = y_{it} - (1 - \sqrt{\sigma^2/(T\sigma_\mu^2 + \sigma^2)})\frac{1}{T} \sum_{i=1}^N y_{it}$

et $x'_{it} = x_{it} - (1 - \sqrt{\sigma^2/(T\sigma_\mu^2 + \sigma^2)})\frac{1}{T} \sum_{i=1}^N x_{it}$.

Fonction de vraisemblance pour le cas d'un modèle SEM

La fonction de vraisemblance dans le cas d'un modèle SEM à effet fixe spatial est comme suit :

$$\begin{aligned} \text{Log}L = & -\frac{NT}{2}\log(2\Pi\sigma^2) + T\log|I_N - \lambda W| - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T [y'_{it} \\ & - \lambda(\sum_{j=1}^N w_{ij}y_{jt})' - (x'_i - \lambda(\sum_{j=1}^N w_{ij}x_{jt})')\beta]^2 \end{aligned} \quad (4.10)$$

Dans le cas où l'effet spatial est considéré comme aléatoire, la fonction de vraisemblance prend la forme suivante :

$$\begin{aligned} \text{Log}L = & -\frac{NT}{2}\log(2\Pi\sigma^2) - \frac{1}{2}\log|V| + (T-1) \sum_{i=1}^N \log|B| \\ & - \frac{1}{2\sigma^2} e' \left(\frac{1}{T} \iota_T \iota_T' \otimes V^{-1} \right) e - \frac{1}{2\sigma^2} e' \left(I_T - \frac{1}{T} \iota_T \iota_T' \right) \otimes (B' B) e, \end{aligned} \quad (4.11)$$

où $V = T\sigma_\mu^2/\sigma^2 I_N + (BB')^{-1}$, $B = I_N - \lambda W$, $e = Y - X\beta$ et ι_T est une colonne de 1 et de taille T. Pour simplifier le calcul numérique du déterminant de la matrice V, Elhorst (2003) propose la relation suivante :

$$\log|V| = \sum_{i=1}^N \log \left[T\sigma_\mu^2/\sigma^2 + \frac{1}{(1 - \lambda\omega_i)^2} \right]. \quad (4.12)$$

4.3.2. Règles de choix des modèles spatiaux

Nous présentons dans cette sous section les différents tests en économétrie spatiale permettant de discriminer entre une spécification autorégressive (SAR ou SDM) et une spécification à autocorrélation des erreurs (SEM).

Dans le cas de données en coupe transversale, la statistique la plus utilisée pour tester l'autocorrélation spatiale est la statistique de Moran (1948, 1950a,b). Comme nous l'avons précisé dans le chapitre précédent, la statistique de Moran est largement utilisée pour une analyse globale (statistique I de Moran globale) et aussi, pour une analyse locale d'autocorrélation (statistique I de Moran locale)⁹.

Concernant la statistique de Moran globale, nous avons montré dans le chapitre précédent¹⁰, qu'elle n'est pas liée à un modèle statistique spécifié. En effet, nous l'avons employée pour établir une analyse exploratoire de nos données d'innovation avant régression. Par ailleurs, cette même statistique peut être utilisée pour tester l'autocorrélation spatiale dans les résidus d'une régression à travers le test de Moran après régression. Néanmoins, ce test n'est utilisé que pour le cas de données en coupe, estimées par MCO.

Pour surmonter cette insuffisance, les tests du multiplicateur de Lagrange (LM) ont été développés (Anselin et al, 1996 ; Anselin et al, 2008 ; Millo et Piras, 2012) afin de tester l'omission d'une autocorrélation spatiale des erreurs ou bien l'omission d'une variable endogène décalée spatialement dans le cadre de données en panel spatial.

4.3.2.1. Les tests du multiplicateur de Lagrange

Depuis les travaux fondateurs de Breusch et Pagan (1980), deux tests de multiplicateurs de Lagrange (LM) ont été développés pour tester l'autocorrélation spatiale pour le cas de données de panel (Baltagi et al., 2005). Récemment, Millo et Piras (2012) ont repris le travail effectué par Baltagi et al. (2005) pour construire ces deux tests. Le choix du modèle spatial SAR ou SEM s'effectue désormais grâce aux tests du multiplicateur de Lagrange suivants : LM_ρ pour le modèle SAR et LM_λ pour le modèle SEM (Anselin, 1988), que nous allons présenter maintenant¹¹.

9. Le I de Moran local ou *Local Indicators of Spatial Association* (LISA) permet de mesurer l'autocorrélation spatiale pour chaque unité spatiale, du fait qu'elle est décomposable à des échelles spatiales assez désagrégées. Voir chapitre 3 pour plus de détail sur les outils de mesure de l'autocorrélation spatiale locale.

10. Voir section 4 du chapitre 3.

11. Voir l'article de Millo et Piras (2012) pour plus de détails sur les tests du multiplicateur de Lagrange pour le cas de données de panel spatial

Le test du multiplicateur de Lagrange LM_ρ :

$$LM_\rho = \frac{[e'(I_T \otimes W)Y/\hat{\sigma}^2]^2}{J}. \quad (4.13)$$

Le test du multiplicateur de Lagrange LM_λ :

$$LM_\lambda = \frac{[e'(I_T \otimes W)e/\hat{\sigma}^2]^2}{T \times T_w}, \quad (4.14)$$

où \otimes est le produit kronecker et e ($e = Y - X\beta$) est le vecteur des résidus du modèle estimé sur données empilées dans les effets spécifiques μ_i et $time_t$.

J est donnée par :

$$J = \frac{1}{\hat{\sigma}^2} [((I_T \otimes W)X\hat{\beta}')(I_{NT} - X(X'X)^{-1}X')(I_T \otimes W)X\hat{\beta} + TT_w\hat{\sigma}^2], \quad (4.15)$$

avec $T_w = tr(WW + W'W)$.

Contrairement aux tests du multiplicateur de Lagrange basés sur les résidus des estimations par la méthode MCO, ces derniers tests sont établis sur les résidus d'une estimation par le maximum de vraisemblance.

Sous l'hypothèse $H_0 : \rho = 0$ ($\lambda = 0$), les statistiques LM_ρ et LM_λ suivent asymptotiquement une loi $N(0, 1)$. Les tests LM_ρ et LM_λ sont les tests robustes les plus utiles dans le choix du modèle spatial approprié (Millo et Piras, 2012). En effet, ces deux tests permettent de vérifier la corrélation spatiale dans les erreurs d'un modèle qui incorpore éventuellement un effet aléatoire.

Après avoir présenté les tests du multiplicateur de Lagrange pour le cas de données de panel spatial, il est nécessaire de choisir entre un modèle à effets aléatoires et un modèle à effets fixes. Pour cela, nous présentons dans ce qui suit le test d'Hausman dans le cas spatial, permettant d'établir un choix entre ces deux types de modèles.

4.3.2.2. Test d'Hausman dans le cas spatial

Avant de passer aux estimations de nos modèles spatiaux, il est nécessaire de tester l'absence ou la présence de corrélation entre les perturbations et la variable explicative du modèle. Généralement, c'est le test d'Hausman (Hausman, 1978) qui compare les effets fixes aux effets aléatoires et teste si oui ou non l'hypothèse de l'effet aléatoire est prise en compte par les données. Tout récemment, Mutl et Pfaffermayr (2011) montrent comment étendre cette procédure dans un cadre spatial (Millo et Piras, 2012). Le test

d'Hausman dans le cadre spatial s'écrit sous cette forme :

$$H = NT(\hat{\theta}_{FGLS} - \hat{\theta}_W)^T (\hat{\Sigma}_W - \hat{\Sigma}_{FGLS})^{-1} (\hat{\theta}_{FGLS} - \hat{\theta}_W), \quad (4.16)$$

où $\hat{\theta}_{FGLS}$ et $\hat{\theta}_W$ sont, respectivement, l'estimateur des moindres carrés généralisés spatiaux et l'estimateur *within*. $\hat{\Sigma}_W$ et $\hat{\Sigma}_{FGLS}$ sont les estimations correspondantes aux coefficients des matrices des variances-covariances. H est asymptotiquement distribuée selon une loi de χ^2 à k degrés de liberté, avec k , le nombre de régresseurs dans le modèle.

Le test d'Hausman appliqué sur nos données en panel spatial¹² rejette l'hypothèse d'absence de corrélation entre les perturbations et les variables explicatives. La valeur du test correspond à 183.35 avec une p -value $< 2.2e^{-16}$. Compte tenu de ces résultats, nous privilégions l'adoption du modèle spatial à effets fixes.

4.4. Résultats des estimations des modèles SAR, SEM et SDM

L'influence significative et positive de la R&D locale sur l'innovation a été largement mise en évidence dans la littérature (Crépon et Duguet, 1997; Licht et Zoz, 1998). Cette relation a été vérifiée par plusieurs modèles et méthodes d'estimation. Par contre, certaines études remettent en cause l'effet de la R&D du voisinage. Nous pensons que la prise en compte de la dimension spatiale pourrait permettre de retrouver le vrai impact de la R&D du voisinage. Lorsque les données sont géoréférencées, l'utilisation des méthodes spatiales permet d'améliorer les résultats d'estimation.

Pour cela, nous exprimons ces relations spatiales à travers une matrice de poids de corrélations spatiales avant d'établir nos régressions. La matrice de poids utilisée pour nos estimation est la matrice de contiguïté d'ordre 1¹³.

Pour mémoire, ces estimations portent sur un échantillon de 94 départements français métropolitains (Corse exclue) couvrant la période 2002-2008. L'estimation se fait par maximum de vraisemblance (MV)¹⁴. Les résultats sont présentés pour les trois modèles spatiaux (SAR, SEM et SDM) sous trois spécifications : spécification avec effets fixes individuels ou spatiaux, spécification avec effets fixes temporels et spécification avec effets fixes individuels et temporels (mixte).

12. La commande *sphstest* du package *splm* du logiciel R est utilisée pour calculer le test d'Hausman dans le cadre spatial. L'estimateur convergent utilisé pour ce cas est l'estimateur Within. Voir l'article de Millo et Piras (2012), pour plus de détails sur ce test.

13. Voir section 3 du chapitre 3 pour une justification du choix d'une matrice de contiguïté d'ordre 1.

14. Comme nous l'avons montré dans la section précédente, les estimateurs MCO ne sont plus convergents en présence d'autocorrélation spatiale.

4.4.1. Tableaux des résultats

Afin d'estimer nos données en panel à effets fixes avec dépendance spatiale au niveau de la variable dépendante (SAR), au niveau du terme d'erreur (SEM) et au niveau des variables explicatives (SDM), nous retenons les spécifications suivantes : modèle spatial à effets fixes individuels, modèle spatial à effets fixes temporels et modèle spatial à effets fixes mixtes.

Les résultats de l'estimation par maximum de vraisemblance sont donnés dans le tableau (4.1), (4.2) et (4.3). Ces estimations incluent moins de variables explicatives que les modèles à effets aléatoires car il est impossible d'estimer les paramètres des variables invariantes dans le temps dans un modèle à effets fixes :

4.4.1.1. Résultat d'estimation du modèle SAR

Tableau 4.1 – Estimation du modèle SAR avec effets fixes

Modèle	Avec effets fixes individuels	Avec effets fixes temporels	Avec effets fixes mixtes
$\ln(R\&Dint_{it-1})$	0.100**** (0.022)	0.646**** (0.027)	0.114**** (0.029)
$\ln(R\&Dextpub_{it-1})$	0.038** (0.010)	0.02* (0.014)	0.018* (0.010)
$\ln(R\&Dextpiv_{it-1})$	0.012 (0.012)	0.021 (0.021)	0.019 (0.013)
$EGindex_{it}$	0.106**** (0.023)	-0.057* (0.035)	0.113**** (0.024)
$EGindex0100_{it}$	0.003 (0.003)	0.0003 (0.067)	0.002 (0.002)
ρ	0.109*	0.297****	0.024*
LR	2.141	89.761	0.089
$LogL$	-34.454	-449.643	-28.004
AIC	270.91	925.29	269.31

LR : Test du rapport de vraisemblance

$LogL$: Logarithme de la vraisemblance

AIC : Critère d'information d'Akaike pour la sélection des modèles.

Les niveaux de significativité statistique sont indiqués par : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

La dépendance spatiale agit probablement aussi par l'intermédiaire de facteurs inobservables. Il est donc probable que cette dépendance spatiale "transite par l'intermédiaire" des perturbations. Pour cela, nous présentons dans le tableau suivant les résultats d'estimations du modèle SEM avec les trois spécifications précédentes :

4.4.1.2. Résultat d'estimation du modèle SEM

Tableau 4.2 – Estimation du modèle SEM avec effets fixes

Modèle	Avec effets fixes	Avec effets fixes	Avec effets fixes
	individuels	temporels	mixtes
$\ln(R\&Dint_{it-1})$	0.114**** (0.022)	0.692**** (0.028)	0.120**** (0.029)
$\ln(R\&Dextpub_{it-1})$	0.022** (0.010)	0.041**** (0.014)	0.018 (0.010)
$\ln(R\&Dextpiv_{it-1})$	0.035*** (0.019)	0.006 (0.022)	0.021 (0.013)
$EGindex_{it}$	0.105**** (0.023)	-0.097**** (0.026)	0.013**** (0.024)
$EGindex0100_{it}$	0.002 (0.003)	0.0005 (0.005)	0.002 (0.002)
λ	0.172**	0.387****	0.098*
LR	4.519	12.275	1.39
$LogL$	-33.265	-488.38	-28.004
AIC	268.53	1002.8	268.01

LR : Test du rapport de vraisemblance

LogL : Le logarithme de la vraisemblance

AIC : Critère d'information d'Akaike pour la sélection des modèles.

Les niveaux de significativité statistique sont indiqués par : * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

La dépendance spatiale agit aussi par l'intermédiaire des variables explicatives. Nous utilisons donc le modèle SDM pour estimer nos données.

4.4.1.3. Résultat d'estimation du modèle SDM

Tableau 4.3 – Estimation du modèle SDM avec effets fixes

Modèle	Avec effets fixes		
	individuels	temporels	mixtes
$\ln(R\&Dint_{it-1})$	0.1**** (0.022)	0.645**** (0.027)	0.120**** (0.029)
$\ln(R\&Dextpub_{it-1})$	0.027** (0.010)	0.037** (0.0142)	0.022** (0.011)
$\ln(R\&Dextpiv_{it-1})$	0.036*** (0.012)	0.102 (0.210)	0.022* (0.013)
$EGindex_{it}$	0.104**** (0.023)	-0.468 (0.354)	0.108**** (0.024)
$W \times \ln(R\&Dint_{it-1})$	-0.017 (0.020)	0.121 (0.078)	-0.016 (0.021)
$W \times \ln(R\&Dextpub_{it-1})$	0.0002 (0.0002)	0.009 (0.032)	0.0002 (0.0002)
$W \times \ln(R\&Dextpiv_{it-1})$	0.0311* (0.0176)	0.0304* (0.0177)	0.030* (0.017)
$W \times EGindex_{it}$	0.013* (0.007)	0.012* (0.007)	0.012* (0.007)
ρ	0.140**	0.320****	0.044*
LR	3.099	15.647	0.299
$LogL$	-27.765	-443.436	-15.79
AIC	267.53	922.87	253.58

LR : Test du rapport de vraisemblance

LogL : Le logarithme de la vraisemblance

AIC : Critère d'information d'Akaike pour la sélection des modèles.

Les niveaux de significativité statistique sont indiqués par : * $p < 0.10$, ** $p < 0.05$,

*** $p < 0.01$, **** $p < 0.001$.

Les questions qui se posent à ce niveau sont les suivantes : lequel des 3 modèles spatiaux est le plus adéquat (§ 4.4.2) ? Une fois le modèle spatial est choisi, laquelle des 3 spécifications (effets fixes individuels, effets fixes temporels et effets fixes mixtes) doit être retenue pour analyser nos résultats (§ 4.4.3) ?

4.4.2. Critères du choix du modèle spatial sur données de panel

Il est primordial de choisir l'un des trois modèles spatiaux (SAR, SEM ou SDM) afin pouvoir interpréter convenablement nos résultats. En effet, nous avons présenté les estimations en utilisant les trois modèles spatiaux SAR, SEM et SDM, mais nous ne savons pas quel modèle domine. Jusqu'à une date très récente, nous ne disposions pas de tests en panel. Comme nous l'avons précisé dans la section précédente, les deux tests du multiplicateur de Lagrange sont appliqués seulement pour des données en coupe (Anselin et Rey, 1991 ; Florax et Folmer, 1992). Tout récemment, Millo et Piras (2012) ont repris le travail de Baltagi et al. (2005) pour construire deux tests du multiplicateur de Lagrange : LM_ρ pour le modèle SAR et LM_λ pour le modèle SEM.

Nous effectuons donc ce test pour pouvoir choisir entre un modèle à variables décalées spatialement (SAR) ou (SDM)¹⁵ et un modèle à erreurs spatialement autocorrélés (SEM).

Tableau 4.4 – Test du Multiplicateur de Lagrange

	Test	p-value
LM_ρ	18.397	0.000
LM_λ	7.129	0.000

Nous remarquons que la valeur du test LM_ρ est largement supérieure à celle du test LM_λ . Nous retenons donc le modèle à variable spatiale décalée. Par ailleurs, la variable "*Spatial Lag*" peut être introduite au niveau de la variable à expliquer et en même temps au niveau des variables explicatives d'un modèle : le modèle SDM. Cette dernière spécification avait reçu beaucoup moins d'attention que les autres, mais son utilisation devient de plus en plus populaire dans les travaux en économétrie spatiale (Baltagi et al., 2005 ; Beer et Riedl, 2012). En effet, le modèle SDM permet d'estimer les *spillovers* provenant de sources différentes (les brevets ou les dépenses en R&D dans notre cas). Pour toutes ces raisons, nous retenons le modèle SDM pour analyser nos résultats¹⁶.

15. Le modèle SDM est aussi un modèle à variable décalée spatialement. Nous pouvons nous limiter au calcul des tests du multiplicateur de Lagrange pour le cas du modèle SAR et du modèle SEM. Si $LM_{\rho(SAR)} > LM_\lambda$ alors $LM_{\rho(SDM)} > LM_\lambda$. Ce test nous permet de choisir entre un modèle SAR et SEM ou entre un modèle SDM et SEM.

16. Ce test nous permet de choisir entre un modèle SEM vs SAR ou bien SEM vs SDM et non pas entre SAR vs SDM. Le modèle SDM contient des variables explicatives décalées spatialement et une variable à expliquer décalée aussi spatialement ; nous retenons donc ce modèle pour interpréter nos résultats. En effet, le modèle SDM nous donne plus d'informations sur les *spillovers* géographiques issus à la fois des brevets et des dépenses en R&D

4.4.3. Analyse des résultats et critères du choix du modèle avec effets fixes adéquat

Après avoir choisi le modèle SDM, notre préoccupation ici est d'examiner la variation des résultats de ses trois spécifications (effets fixes individuels, effets fixes temporels et effets fixes mixtes). Nous remarquons que les résultats varient selon la spécification appliquée. Si nous considérons le modèle avec effets fixes temporels, nous remarquons que l'effet de la R&D interne est élevé (0.645). Néanmoins, cet effet devient relativement faible si la spécification est avec effets fixes individuels ou avec effets fixes mixtes (respectivement 0.100 et 0.120). De même, l'indice d'Ellison Glaeser indique que la diversification industrielle est bénéfique seulement pour les cas des modèles avec effets fixes individuels et avec effets mixtes. Cette variation au niveau des coefficients montre que la prise en compte de l'hétérogénéité individuelle entre les observations (94 départements dans notre cas) via μ_i , a un effet sur l'ampleur des coefficients estimés. De plus, les estimations du modèle autorégressif avec effets fixes temporels, individuels ou mixtes montrent que le coefficient ρ est significatif : l'innovation (logarithme des brevets) des départements voisins à un département i influence l'innovation départementale locale.

Par ailleurs, le coefficient ρ est plus important pour le cas du modèle avec effets fixes temporels (0.320 contre 0.140 pour un modèle avec effets fixes individuels et 0.044 pour un modèle avec effets mixtes). Ceci est expliqué par le fait que pour une spécification avec effets fixes temporels, les individus, qui représentent les départements dans notre cas et qui sont eux mêmes des variables spatiales, ne sont pas pris en compte dans l'estimation. Ainsi, ceci explique l'ampleur du coefficient spatial ρ pour le cas d'un modèle avec effets fixes temporels. Ça explique aussi l'ampleur du coefficient de la R&D interne.

Même si ce dernier modèle ne semble pas pertinent pour estimer nos données d'innovation, il est important de se baser sur un test statistique afin de choisir l'une de ces trois spécifications¹⁷.

Pour cette fin, nous utilisons le critère d'information d'Akaike (AIC) le plus bas¹⁸ (Pace et LeSage, 2004) et nous choisissons ainsi la troisième spécification¹⁹.

17. Nous pouvons utiliser notamment le test du rapport de vraisemblance LR pour comparer les trois modèles. Le test LR confirme notre choix puisque c'est le modèle avec effets fixes mixtes qui a la valeur LR la plus faible.

18. Ou bien le Log de vraisemblance le plus élevé.

19. Les résultats trouvés dans le tableau (4.2), avec un modèle SEM, sont similaires à ceux trouvés avec un modèle SAR. Les effets de la R&D interne sont toujours supérieurs à ceux de la R&D externe qu'il s'agisse d'un modèle à effets fixes individuels, à effets fixes temporels ou à effets fixes mixtes. Ainsi, quelle que soit la spécification et quels que soient les effets, l'estimation montre que le coefficient

Généralement, le modèle avec effets fixes mixtes (two ways fixed model)²⁰ est le plus adopté dans la mesure où il influence positivement les coefficients estimés (Baltagi, 2005). En conclusion, nous choisissons la troisième spécification (modèle à effets fixes mixtes) pour analyser nos résultats.

Les résultats d'estimation de la FPI avec le modèle SDM à double effets fixes sont très proches de ceux trouvés avec des estimations en panel sans prise en compte de l'autocorrélation spatiale (décrits dans le chapitre 2 de notre travail). Les résultats indiquent un impact significatif des dépenses en R&D sur le logarithme du nombre de brevets affectés aux départements. Les élasticités estimées se situent entre 0.022 (pour $\ln(R\&Dextpub_{it-1})$) et 0.120 (pour $\ln(R\&Dint_{it-1})$). Ces régressions donnent un coefficient légèrement inférieur pour les dépenses en R&D interne (0.120 au lieu de 0.13 (voir tableau 2.6) et des coefficients presque identiques pour les dépenses en R&D externe publique et privée.

Les résultats du modèle SDM confirment *la proposition 1*²¹ formulée dans le chapitre 2 puisque le coefficient des dépenses en R&D interne est plus élevé que les deux coefficients des dépenses en R&D externe²². De plus, l'effet de la R&D externe privée des départements voisins ($W \times \ln(R\&Dextpiv)_{it-1}$) est positif sur les activités innovantes du département considéré. Le coefficient de la diversité régionale d'Ellison-Glaeser ($EGindex_{it}$) est légèrement plus élevé dans les estimations spatiales en panel. De même, le Ellison-Glaeser des régions voisines ($W \times EGindex)_{it}$ semble avoir un impact significatif et positif sur les activités départementales des brevets dans le cadre d'un modèle SDM.

Par ailleurs, les coefficients estimés de ce modèle méritent une attention particulière (Anselin et Le Gallo, 2006; Kelejian et al., 2006; LeSage et Autant-Bernard, 2011). En effet, l'interprétation des coefficients dans un modèle spatial autorégressif est problématique dans la mesure où les effets marginaux ne correspondent pas aux coefficients estimés. Ainsi, une correction a été proposée par LeSage et Pace (2009). Elle consiste à décomposer l'effet des variables explicatives en effet direct et en effet indirect afin de rendre l'interprétation des coefficients de ce modèle spatial plus éloquente.

autorégressif spatial est toujours significatif : une dépendance spatiale agit par l'intermédiaire de facteurs inobservables entre départements. Elle "transite" par l'intermédiaire des perturbations entre voisins géographiques.

20. Un autre argument en défaveur du modèle à effets fixes temporels est que 2002-2008 est une période sans choc conjoncturel sur la R&D ; la récession a eu lieu en 2009 en France.

21. *Proposition 1* : La nature des connaissances a un impact significatif sur l'ampleur des externalités de connaissances. Par conséquent, la R&D interne a un effet plus élevé sur l'innovation que la R&D externe puisque cette dernière ne mesure que les connaissances codifiées sans tenir compte des connaissances tacites.

22. Nous avons aussi vérifié dans des estimations non présentées ici que cela n'est pas du à la distinction R&D externe privée/ R&D externe publique.

Nous allons préciser dans ce qui suit l'importance de cette interprétation dans le cas du modèle SDM dans la mesure où nous pourrions interpréter à la fois le décalage spatial au niveau de la variable dépendante et au niveau des variables explicatives. L'interprétation des paramètres deviendra donc plus riche (LeSage et Pace, 2009) si nous décomposons ces effets en effets directs et indirects²³.

4.5. Décomposition de l'effet spatial en effets directs et effets indirects

L'étude des trajectoires des externalités de connaissances est un thème qui a été traité depuis longtemps. Les modèles utilisés pour cette fin proposent de quantifier l'ampleur des *spillovers* de nature spatiale. Généralement, il s'agit de mesurer les *spillovers* résultant de l'innovation technologique des régions voisines d'une région quelconque, et qui ont un effet positif sur la production de l'innovation au sein de cette région (LeSage et Pace, 2009). En effet, les externalités de connaissances sont considérées de nature locale et relevant seulement des régions voisines ce qui s'explique par le fait que la connaissance tacite liée à l'expérience des inventeurs et des chercheurs "*ne voyage pas bien*" (LeSage et Pace, 2009). Ce stock de connaissances augmente dans la région où les inventeurs locaux découvrent de nouvelles idées et les diffusent à travers les contacts face à face.

En utilisant le modèle SDM pour estimer nos données d'innovation, nous pouvons mesurer les externalités de connaissance produites par les brevets et aussi celles produites par les dépenses en R&D à travers les coefficients suivants :

- $\hat{\rho} \sum_{j=1}^N w_{ij} y_{jt}$ permet de mesurer les externalités de connaissance générées par la variable à expliquer (Anselin et al., 1997, Acs et al., 2002).
- $\sum_{j=1}^N w_{ij} x_{jt} \theta$ permet de mesurer les externalités de connaissance générées par les variables explicatives (LeSage et Pace, 2009).

Ces variables sont introduites dans le modèle afin de rendre compte de l'existence des externalités de connaissance. En effet, les unités géographiques bénéficient davantage des externalités lorsqu'elles sont situées dans un environnement riche et fertile qui se manifeste à travers un nombre élevé de demandes de brevets et des dépenses importantes en R&D (externalités générées par les variables explicatives du modèle). L'estimation de ces paramètres est une source d'informations sur la relation entre observations. En effet, la capacité des modèles de régression spatiaux à saisir ces interactions représente un apport important (Behrens et Thisse, 2007).

23. Contrairement au modèle SDM, le modèle SEM ne présente pas de décalage spatial au niveau de la variable dépendante et donc nous ne pouvons pas parler d'effets indirects découlant des changements dans les variables explicatives.

Cependant, ce riche ensemble d'informations augmente également la difficulté de l'interprétation des estimations qui en résultent. Compte tenu de cette difficulté dans l'interprétation des coefficients, il serait intéressant de décomposer l'effet spatial en effet direct et en effet indirect pour montrer l'effet des externalités de connaissance (représentées généralement par les effets indirects) dans le processus de l'innovation des départements voisins. Autrement dit, le changement d'une seule observation (département) associée à une variable explicative donnée aura une incidence sur la région elle-même (effet direct) et est susceptible d'affecter toutes les autres départements indirectement, puis d'affecter le département initial en retour (effet indirects transitant par les départements voisins).

Cette section est destinée à présenter l'importance de la décomposition du modèle SDM estimé sur données de panel avec effets fixes mixtes, en effets directs et effets indirects (§ 4.5.1), avant de présenter les résultats des estimations (§ 4.5.2).

4.5.1. Du modèle SDM aux effets directs et indirects

Dans un modèle linéaire simple de la forme $y = \sum_{r=1}^k x_r \beta_r + \epsilon$ où les observations sont supposées être indépendantes, β_r n'est autre que la dérivée partielle de y_i par rapport à x_{ir} (effet direct) : $\frac{\partial y_i}{\partial x_{ir}} = \beta_r$ pour tout i et $\frac{\partial y_i}{\partial x_{jr}} = 0$ pour $j \neq i$.

Autrement dit, une modification de la variable explicative sur l'observation i n'entraîne de modifications que sur elle-même et non sur les autres observations.

Dans le cas d'un modèle SDM, les variables explicatives de la région i exercent, non seulement un effet direct sur la variable dépendante de la même région, y_i , mais aussi un effet indirect sur celles des régions voisines, y_j pour ($j \neq i$).

Ce second type d'effet indirect est généré du fait que les observations sont considérées comme dépendantes spatialement.

Pour détecter l'effet indirect, LeSage et Pace (2009) présentent le modèle dans le cas d'une étude en coupe transversale, sous la forme matricielle suivante :

$$y(I_n - \rho W) = X\beta + WX\theta + l_n\alpha + \epsilon \quad (4.17)$$

$$y = \sum_{r=1}^k S_r(W)x_r + V(W)l_n\alpha + V(W)\epsilon \quad (4.18)$$

$$S_r(W) = V(W)(I_n\beta_r + W\theta_r) \quad (4.19)$$

$$V(W) = (I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots \quad (4.20)$$

Encadré 4.2 : Définition de $S_r(W)$ (équation 4.19)

Pour comprendre la signification de $S_r(W)$, reprenons cette équation :

$$y = \sum_{r=1}^k S_r(W)x_r + V(W)l_n\alpha + V(W)\epsilon \quad (4.21)$$

Ceci peut être présenté comme suit :

$$\begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ y_n \end{pmatrix} = \sum_{r=1}^k \begin{pmatrix} S_r(W)_{11} & \dots & S_r(W)_{1n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ S_r(W)_{n1} & \dots & S_r(W)_{nn} \end{pmatrix} \times \begin{pmatrix} x_{1r} \\ \cdot \\ \cdot \\ x_{nr} \end{pmatrix} + V(W)l_n\alpha + V(W)\epsilon \quad (4.22)$$

Nous utilisons $S_r(W)_{ij}$ pour dénoter le $i^{\text{ème}}$ et le $j^{\text{ème}}$ élément de la matrice $S_r(W)$ et $V(W)_i$ pour noter la $i^{\text{ème}}$ ligne de $V(W)$.

$$y_i = \sum_{r=1}^k [S_r(W)_{i1}x_{1r} + S_r(W)_{i2}x_{2r} + \dots + S_r(W)_{in}x_{nr}] + V(W)_i l_n \alpha + V(W)_i \epsilon. \quad (4.23)$$

Contrairement au modèle linéaire simple qui suppose l'indépendance entre les observations, l'équation (4.19) montre bien que :

$$\frac{\partial y_i}{\partial x_{jr}} = S_r(W)_{ij} \neq 0 \quad (4.24)$$

$S_r(W)$ représente l'effet de changement d'une observation d'une variable explicative sur la variable dépendante. L'effet des autres départements est pris en compte dans le modèle SDM à travers l'introduction des variables WY et WX . Ce résultat est très important dans l'interprétation des coefficients de la fonction de production d'innovation. En effet, les variations départementales des niveaux d'innovation dépendent des niveaux d'innovation des départements voisins capturés par le vecteur de décalage spatial WY , ainsi que par les dépenses en R&D des régions voisines représentées par WX .

L'impact d'un changement au niveau de la variable explicative x_{ir} sur la variable dépendante y_i est le suivant :

$$\frac{\partial y_i}{\partial x_{ir}} = S_r(W)_{ii} \quad (4.25)$$

Cet impact (effet direct) tient compte du fait que le changement produit sur x_{ir} a aussi un impact sur le voisinage de i (les y_j tels que j est voisin de i), ce qui contribue de nouveau à impacter y_i (LeSage et Pace, 2009). L'effet direct est présenté par les éléments de la diagonale principale de la matrice $S_r(W)$ alors que les éléments hors diagonale présentent l'effet indirect. Cet effet indirect provient du fait que les brevets du département i influencent les brevets des départements voisins, qui influencent en retour les brevets du département i via le coefficient ρ . Ce dernier représente l'effet des *spillovers* (LeSage et Pace, 2009). Ces impacts directs et indirects sont généralement utilisés pour étudier plusieurs phénomènes. Par exemple, Kelejian et al. (2006) examinent l'impact de la contagion financière d'un pays sur d'autres pays, Anselin et Le Gallo (2006) examinent la diffusion de la pollution de l'air, et Baumont et al. (2003) et Guillain et Le Gallo, (2007) examinent les incidences de l'évolution des fonds structurels européens dans des régions stratégiques sur la croissance économique globale.

En se basant sur les équations (4.23) et (4.24), LeSage et Pace (2009) suggèrent une mesure synthétique de ces impacts. Étant donné que les effets directs (représentés par les éléments de la diagonale principale de la matrice $S_r(W)$) et les effets indirects (représentés par les éléments hors diagonale) varient d'une région à une autre, LeSage et Pace (2009) considèrent la moyenne de chacun de ces effets comme étant respectivement les effets directs et les effets indirects de chaque région. La somme (sur les lignes ou sur les colonnes) des effets directs ou des effets indirects n'est autre que l'effet total :

- Une moyenne de la diagonale de la matrice $S_r(W)$ fournit une mesure synthétique de l'impact direct moyen.
- La différence entre l'impact total moyen et l'impact direct moyen constitue l'impact indirect moyen.
- L'impact total moyen est la somme des impacts totaux sur les lignes (ou colonnes) de la matrice $S_r(W)$, puis divisé sur le nombre des régions.

Formellement, les définitions de ces mesures sont les suivantes (LeSage and Pace, 2009 ; p. 38 (équation 17)) :

- *Impact direct moyen* : l'impact du changement de la $i^{\text{ème}}$ observation de x_{ir} sur y_i peut être mesuré en utilisant $S_r(W)_{ii}$, et est égale à $n^{-1}tr(S_r(W))$. L'impact direct associé à toutes les observations i est similaire aux coefficients d'une régression qui représentent une réponse de la variable dépendante à toutes les autres variables indépendantes.
- *Impact total moyen sur une observation* : la somme de la $i^{\text{ème}}$ ligne représente l'impact total sur une observation individuelle y_i résultante du changement de la r^{th} variable explicative sur toutes les autres observations (régions).

- Impact total moyen d'une observation sur les autres observations : c'est l'impact de changement d'une observation i sur toutes les autres observations.

Les formules finales de ces effets sont représentées comme suit :

$$\bar{M}(r)direct = n^{-1}tr(S_r(W))$$

$$\bar{M}(r)total = n^{-1}l'_n S_r(W) l_n$$

$$\bar{M}(r)indirect = \bar{M}(r)total - \bar{M}(r)direct$$

avec tr désigne la trace de la matrice.

4.5.2. Effets *feedbacks* ou effets de rétroaction simultanée

Dans cette section, nous montrons comment l'effet direct peut être décomposé en deux effets : un effet sur la variable dépendante et un effet réciproque (*feedback loops*) exercé par la région voisine j sur la région i elle-même (LeSage et Pace, 2009).

L'effet *feedback* est une caractéristique du modèle de régression spatiale résultant de la relation de dépendance spatiale entre observations. Pour comprendre d'où vient cette différence, nous utilisons l'expression suivante bien connue pour exprimer l'inverse comme une série infinie de Debreu et Herstein (1953) :

$$(I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots,$$

Ceci conduit à une reformulation du modèle spatial autorégressif y en utilisant l'expression autorégressive ci-dessus :

$$y = (I_n - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \epsilon \text{ peut s'écrire donc de cette manière :}$$

$$y = (I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots) X\beta + (I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots) \epsilon \text{ et donc,}$$

$$y = X\beta + \rho W X\beta + \rho^2 W^2 X\beta + \rho^3 W^3 X\beta + \dots + \epsilon + \rho W \epsilon + \rho^2 W^2 \epsilon + \rho^3 W^3 \epsilon + \dots \quad (4.26)$$

avec : $W^1 = W$ est la matrice de contiguïté d'ordre 1 et définie telle que la diagonale soit nulle (une observation ne peut être le voisin de soi-même). W^2 est la matrice de contiguïté d'ordre 2 et représente formellement les observations qui sont voisines des premiers voisins. Dans ce cas, la diagonale W^2 est différente de 0 puisqu'un département i peut être le voisin de ses voisins. De même, la diagonale de W^3 est non nulle elle aussi car il est possible que le département i soit voisin du voisin de ses voisins, etc. L'ampleur de ce "feedback" dépend de :

- la position des départements dans l'espace,
- du degré de connectivité entre les départements qui dépend de la matrice de poids W dans le modèle,
- du paramètre ρ qui mesure la force de la dépendance spatiale,
- des paramètres β et θ .

L'étude de l'impact de changement d'un seul département ou de l'impact des changements sur un seul département serait représenté par un changement d'une colonne ou d'une ligne de la matrice $S_r(W)$. En effet, un département i est considéré comme le voisin de ses voisins de sorte que l'impact sur ces départements voisins exercera une influence sur le département i lui-même : l'effet "*feedback*" où l'observation i affecte l'observation j et l'observation j affecte l'observation i en retour. Il est évident que les effets résultant d'un changement dans la variable explicative vont influencer davantage les voisins les plus proches que les voisins les plus éloignés, et ceci en fonction de l'ordre du voisinage.

L'effet direct est composé donc de deux effets : effet sur la variable dépendante de cette même région, et d'un effet "*feedback effect*" associé aux *spillovers* géographiques.

4.5.3. Résultats de la décomposition en effets directs et indirects

Nous retenons le modèle SDM à effets fixes mixtes (Colonne 3 du tableau 4.3) et nous examinons dans ce qui suit la décomposition de l'effet spatial des coefficients de nos variables en effets directs et en effets indirects.

4.5.3.1. Effets directs et indirects des variables explicatives

Tableau 4.5 – Estimation des effets directs, indirects et totaux du modèle SDM (Effets mixtes du tableau 4.3)

Modèle	effets directs (1)	effets indirects (2)	effets totaux (3)	SDM (3)/(4) (4)	(5)
$\ln(R\&Dint_{it-1})$	0.137	0.034	0.171	0.120	1.425
$\ln(R\&Dextpub_{it-1})$	0.0274	0.0076	0.035	0.022	1.590
$\ln(R\&Dextpriv_{it-1})$	0.045	0.013	0.057	0.022	2.590
$EGindex_{it}$	0.109	0.026	0.135	0.108	1.25

(4) représente les coefficients du modèle SDM à effets fixes mixtes estimés dans le tableau (4.3).

(5) représente le rapport entre les effets totaux (3) et les coefficients du modèle SDM à effets fixes mixtes (4).

La première ligne du tableau (4.5) montre que l'effet direct de la R&D interne est de l'ordre de 0.137. Cet effet englobe l'effet des dépenses en R&D interne du département i sur les brevets du département lui-même et les effets de "*feedbacks*" générés à travers l'augmentation du nombre de brevets des départements voisins ($W^1+W^2+\dots$). Ceci explique le fait que l'ampleur de ce coefficient (0.137) est supérieure au coefficient de la R&D interne (0.120) dans le modèle spatial SDM à effets fixes mixtes (colonne 3 du

tableau (4.3) et colonne (4) du tableau (4.5)). En effet, ce n'est pas seulement l'effet de la R&D interne du département i sur la variable y_i qui est mesuré, mais aussi l'effet de l'augmentation de l'innovation des départements voisins sur la variable y_i , résultante de l'investissement du département i en R&D interne. L'effet "feedback" est de l'ordre de 0.017 (0.137-0.120) dans ce cas.

Concernant l'effet indirect des dépenses en R&D interne (0.034), nous remarquons qu'il est relativement faible (en comparaison avec l'effet direct). Néanmoins, l'effet des *spillovers* est non négligeable. Cet effet est représenté par la somme des effets indirects (0.034) et des "feedbacks" (0.017), ce qui représente un total de *spillovers* de 0.051 d'effets en plus.

Les effets totaux (3) représentent pour chaque variable la somme des effets directs et des effets indirects.

Les variables $\ln(R\&Dextpub_{it-1})$, $\ln(R\&Dextpiv_{it-1})$ et $EGindex_{it}$ présentent respectivement les coefficients directs de valeur 0.0274, 0.045 et 0.109. La différence entre l'effet direct et le coefficient estimé du modèle SDM est plus importante pour le cas de la variable $\ln(R\&Dextpiv_{it-1})$ (0.022).

Comme nous l'avons défini dans le chapitre 2 de notre travail, les dépenses en R&D externes privées intègrent des connaissances codifiées qui circulent mieux dans l'espace que les connaissances tacites. Les dépenses en R&D externe privée d'un département i vont améliorer la capacité productive de ce département ainsi que celles des départements voisins. Nous pouvons donc dire qu'il existe un effet de *spillovers* positif pour ces variables explicatives et surtout pour la variable $\ln(R\&Dextpiv_{it-1})$.

Afin de mieux comprendre la source des effets indirects et des effets "feedbacks", nous proposons dans ce qui suit de répartir en fonction de l'ordre de contiguïté les effets (directs et indirects) de chacune de nos variables explicatives : $\ln(R\&Dint_{it-1})$, $\ln(R\&Dextpub_{it-1})$, $\ln(R\&Dextpiv_{it-1})$ et $EGindex_{it}$, respectivement dans les tableaux (4.6), (4.7), (4.8) et (4.9).

4.5.3.2. Répartition des effets directs et indirects par matrice de contiguïté d'ordre 0 à 4

Le tableau (4.5) nous montre la répartition de l'effet total de nos variables explicatives en effets directs et indirects. Par ailleurs, en utilisant l'équation (4.26), nous pouvons faire une décomposition de ces effets en fonction de la matrice de contiguïté d'ordre 0 à 4²⁴. Cette décomposition permet de mettre en avant un phénomène de diffusion des effets. Dans un premier temps, la modification d'une variable explicative

24. Au delà de l'ordre 4, les effets de *spillovers* deviennent très négligeables.

4.5. Décomposition de l'effet spatial en effets directs et effets indirects

va entraîner des variations de la variable à expliquer au niveau local. Dans un second temps, cela va entraîner une modification sur le plus proche voisinage. Ensuite, ce phénomène va continuer à se propager en entraînant des modifications aussi bien au niveau local que sur le voisinage qui devient de plus en plus éloigné.

Pour cela, nous étudions dans les tableaux suivants, la décomposition des effets directs et indirects pour chacune des variables $\ln(R\&Dint_{it-1})$, $\ln(R\&Dextpub_{it-1})$, $\ln(R\&Dextpiv_{it-1})$ et $EGindex_{it}$, en fonction de l'ordre (de 0 à 4) de la matrice de contiguïté.

Le tableau (4.6) représente une décomposition de la première ligne du tableau (4.5) :

Tableau 4.6 – $\ln(R\&Dint_{it-1})$: Répartition de l'effet direct, indirect et total par matrice de contiguïté d'ordre 0 à 4

Ordre de W	effets directs	effets indirects	effets totaux
W_0	0.120	0.000	0.1357
W_1	0.000	0.027	0.027
W_2	$1.830 \cdot 10^{-04}$	0.005	0.0057
W_3	$1.936 \cdot 10^{-05}$	0.001	0.0011
W_4	$3.802 \cdot 10^{-06}$	0.0002	0.0002
<i>ligne 1 du tab (4.5)</i>	$\Sigma = 0.137$	$\Sigma = 0.034$	$\Sigma = 0.171$

Le tableau (4.7) représente une décomposition de la deuxième ligne du tableau (4.5) :

Tableau 4.7 – $\ln(R\&Dextpub_{it-1})$: Répartition de l'effet direct, indirect et total par matrice de contiguïté d'ordre 0 à 4

Ordre de W	effets directs	effets indirects	effets totaux
W_0	0.022	0.000	0.0271
W_1	0.000	5.57910^{-03}	5.57910^{-03}
W_2	$3.665 \cdot 10^{-05}$	1.10810^{-03}	1.14510^{-03}
W_3	$3.878 \cdot 10^{-06}$	2.31210^{-04}	2.35110^{-04}
W_4	$7.614 \cdot 10^{-07}$	4.75110^{-05}	4.82810^{-05}
<i>ligne 2 du tab (4.5)</i>	$\Sigma = 0.0274$	$\Sigma = 0.0076$	$\Sigma = 0.035$

Le tableau (4.8) représente une décomposition de la troisième ligne du tableau (4.5) :

Tableau 4.8 – $\ln(R\&Dextpiv_{it-1})$: Répartition de l'effet direct, indirect et total par matrice de contiguïté d'ordre 0 à 4

Ordre de W	effets directs	effets indirects	effets totaux
W_0	0.022	0.000	0.045
W_1	0.000	9.31510^{-03}	9.31610^{-03}
W_2	6.12010^{-05}	1.85110^{-03}	1.91210^{-03}
W_3	6.47410^{-06}	3.86110^{-04}	3.92610^{-04}
W_4	$1.271 \cdot 10^{-06}$	7.93410^{-05}	8.06110^{-05}
<i>ligne 3 du tab (4.5)</i>	$\Sigma = 0.045$	$\Sigma = 0.013$	$\Sigma = 0.057$

Le tableau (4.9) représente une décomposition de la quatrième ligne du tableau (4.5) :

Tableau 4.9 – $EGindex_{it}$: Répartition de l'effet direct, indirect et total par matrice de contiguïté d'ordre 0 à 4

Ordre de W	effets directs	effets indirects	effets totaux
W_0	0.1089	0.000	0.1089
W_1	0.000	0.0214	0.0214
W_2	1.35410^{-04}	0.004	0.004
W_3	1.37610^{-05}	0.0008	0.0008
W_4	$2.594 \cdot 10^{-06}$	0.0001	0.0001
<i>ligne 4 du tab (4.5)</i>	$\Sigma = 0.109$	$\Sigma = 0.026$	$\Sigma = 0.135$

Nous avons présenté dans les tableaux (4.6, 4.7, 4.8 et 4.9) la répartition des effets directs et indirects de nos variables explicatives par matrice de contiguïté d'ordre 0 à 4. Nous avons choisi de décomposer ces effets jusqu'à la matrice de contiguïté d'ordre 4 puisque au delà de cet ordre du voisinage, les effets deviennent nuls. La somme de tous les effets directs (ou indirects) de chaque variable n'est autre que son effet direct (ou indirect) calculé dans le tableau (4.5). Par exemple la somme des effets directs de la variable $\ln(R\&Dint_{it-1})$ est égale à 0.137 (tableau 4.5), représenté dans le tableau (4.6) par $0.120+0.000+1.830 \cdot 10^{-04}+1.936 \cdot 10^{-05}+3.802 \cdot 10^{-06}+\dots=0.137$ ²⁵.

25. Même principe pour les effets indirects : la somme des effets indirects de la R&D interne du tableau (4.6) est mesurée par $0.000+0.027+0.005+0.001+0.0002+\dots=0.034$ (effet indirect de la R&D interne du tableau (4.5)).

Pour comprendre l'effet de l'espace sur l'innovation d'un département considéré, nous représentons dans le schéma suivant les effets directs, les effets "feedbacks" et les effets indirects du tableau (4.6)²⁶ :

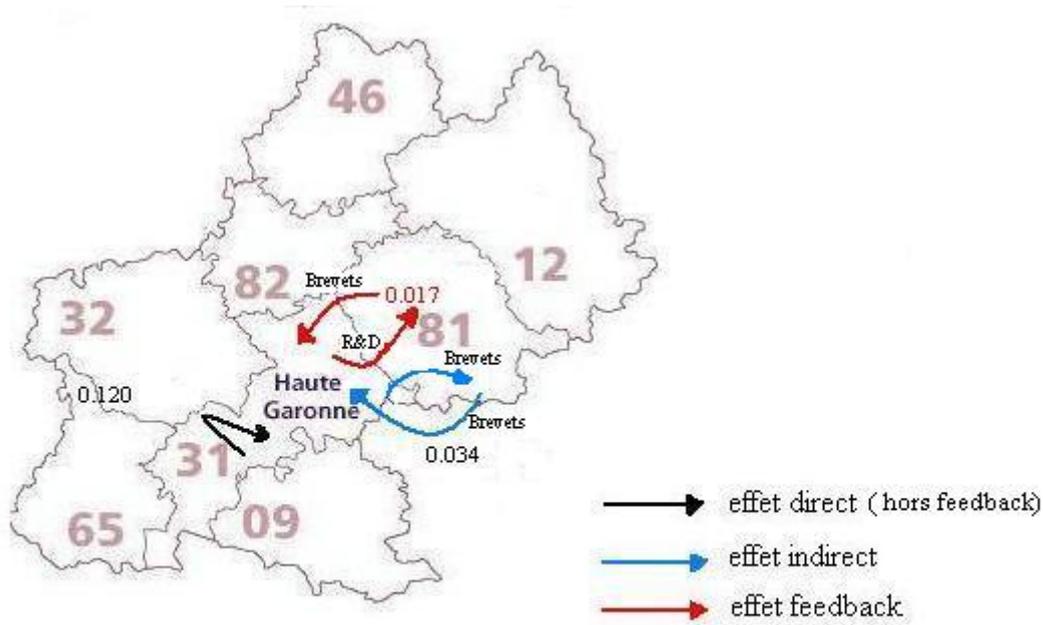


Figure 4.1 – Répartition des effets directs et indirects de la R&D interne entre départements de la région Midi-Pyrénées

A partir du tableau (4.6) et de la figure (4.1), nous essayons d'expliquer les interactions spatiales qui interviennent entre différents départements et comment les dépenses en R&D interne²⁷ des départements voisins peuvent influencer l'innovation d'un département considéré.

Nous prenons le cas du département Haute-Garonne et nous étudions ses interactions avec ses départements voisins.

Trois étapes sont considérées pour ce processus de diffusion des connaissances :

- 1^{ère} étape : La première ligne du tableau (4.6) montre que si la R&D interne du département Haute-Garonne progressait de 1%, l'innovation de ce même département augmenterait de 0.120% (flèche en noir sur la figure 4.1). Comme

26. Pour ne pas surcharger la figure, nous n'avons représenté dans ce schéma qu'un seul département voisin (le Tarn (81)). Il est bien entendu que tous les autres départements voisins sont impliqués dans ce phénomène.

27. Pour simplifier, nous nous focalisons dans cet exemple sur l'étude des interactions entre départements, résultantes des externalités de la R&D interne. Le même principe est appliqué pour les dépenses en R&D externe privée et publique.

l'effet de la R&D interne est sur le département lui-même, la matrice de contiguïté est donc d'ordre 0. Ainsi, l'effet indirect mesurant les *spillovers* (LeSage et Pace, 2009) est nul²⁸.

- 2^{ème} étape : L'innovation des départements voisins au département Haute-Garonne a un effet positif sur l'augmentation de l'innovation du département Haute-Garonne. Ainsi, si les brevets des départements voisins progressaient de 1%, l'innovation (mesurée par le log(brevets)) du département Haute-Garonne augmenterait de 0.034% en moyenne (flèche en bleu).
- 3^{ème} étape : L'investissement en R&D interne du département Haute-Garonne influence aussi l'innovation des départements voisins, notamment les voisins les plus proches (W_1). Une fois que les départements voisins connaissent une augmentation du niveau d'innovation (grâce aux externalités de connaissance produites par le département Haute-Garonne), la production d'innovation du département Haute-Garonne augmente (augmentation de 0.017%)(flèche en rouge).

Ainsi, le département Haute-Garonne reçoit des externalités de connaissance du département Tarn (81) et des autres départements voisins qui, à l'origine, sont le résultat de l'investissement du département Haute-Garonne en R&D interne. Ces effets représentent les effets "*feedbacks*" ou de rétroaction.

A travers la figure (4.1), nous avons montré comment se répartissent les effets directs et indirects selon l'ordre du voisinage. Nous pouvons aussi remarquer que dans les tableaux (4.6, 4.7, 4.8 et 4.9) les effets directs sont une fonction décroissante du nombre de voisins tandis que les effets indirects décroissent plus lentement que les effets directs, au fur et à mesure que nous augmentons le nombre de voisins. Les tableaux indiquent que l'impact direct tend vers 0 au fur et à mesure que nous allons plus loin dans l'espace et que l'effet de dégradation dans l'espace est extrêmement rapide.

En plus, les impacts indirects sur les voisins du second ordre (W_1) passent au 1/5 pour les variables $\ln(R\&D_{int_{it-1}})$, $\ln(R\&D_{extpub_{it-1}})$ par rapport à ceux du voisin d'ordre 1, au 1/9 pour $\ln(R\&D_{extpiv_{it-1}})$ et au 1/18 pour l'indice d'Ellison Glaeser. Cette baisse montre que l'ampleur des *spillovers* de connaissance diminue drastiquement avec la distance. Ceci ressemble aux résultats trouvés par Anselin et al.(1997); Autant-Bernard (2001a); Bottazzi et Peri (2003) et récemment à ceux de LeSage et Autant-Bernard (2011). Intuitivement, il est plus facile de bénéficier de la R&D des départements contigus que de la R&D d'autres départements plus éloignés puisque la connaissance dans ce cas est plus facilement assimilée.

28. Pour W^0 , $S_r W = I_n \beta_r$ avec β est le coefficient de régression des dépenses en R&D interne du modèle SDM.

4.6. Conclusion

L'objectif de ce chapitre était d'examiner la façon dont la dépendance spatiale peut influencer la relation entre dépenses en R&D et innovation. Nous avons mené cette analyse sur les 94 départements français métropolitains (Corse exclue) sur des données en panel pour la période 2002-2008. Le recours aux outils de l'économétrie spatiale sur des données en panel nous a permis de tenir compte d'une manière explicite des effets spatiaux.

Les résultats trouvés dans ce chapitre nous ont permis de montrer que les techniques de modélisation antérieures de la fonction de production de connaissances sont mal spécifiées surtout quand il s'agit de la mesure de l'effet spatial. En effet, les études précédentes s'appuient généralement sur des données en coupe et ne prennent pas en compte l'effet de l'hétérogénéité entre observations. En plus, l'effet *spillovers* a été généralement mesuré par le coefficient $\hat{\rho}Wy$ (Anselin et al., 1997 ; Acs et al., 2002).

A travers la décomposition des coefficients des variables en effets directs et indirects, nous avons montré que les *spillovers* peuvent provenir aussi d'un effet "feedback" réciproque exercé par le département j sur le département i lui-même (LeSage et Pace, 2009).

La décomposition en effets directs et indirects a été établie avec le modèle de Durbin spatial à double effets, SDM en utilisant une matrice de contiguïté d'ordre 1. Le choix de ce modèle a été effectué sur la base des tests robustes du multiplicateur de Lagrange, dédiés aux données en panel spatial (Millo et Piras, 2012). Les résultats d'estimation de ce modèle montrent que les élasticités des dépenses en R&D (interne et externe) présentent des signes attendus quelque soit la spécification retenue (effets fixes, effets individuels ou effets mixtes).

Nos résultats ont montré aussi que quelque soit le modèle estimé (SAR, SEM et SDM), l'effet spatial existe et est d'une ampleur non négligeable.

La R&D interne, qui est le déterminant essentiel de l'innovation mesurée par le logarithme des brevets, a un effet important sur la production d'innovation du département lui-même (0.120). Néanmoins, cet effet ne tient pas compte de l'effet de *feedback* (de l'ordre de 0.017 pour la R&D interne) obtenu grâce à l'innovation des départements voisins au département i , résultant de l'investissement de ce département i en R&D interne. Ainsi la somme de ces deux effets (0.120+0.017) constitue l'effet direct de la R&D interne sur la production d'innovation d'un département i .

A cet effet direct, se rajoute un autre effet : l'effet indirect, ou ce que LeSage et Pace (2009) appellent effet de *spillovers* géographiques. Au total, les externalités de

connaissance capturées dans les effets feedbacks (+0.017) et indirects (+0.034) ajoutent 0.051% à l'effet initial de 0.12% sur les brevets. Ce qui donne un effet total de 0.171 au lieu de 0.120 (donc multiplié par 1.425). Ceci constitue donc l'effet total de la R&D interne sur la production d'innovation d'un département i . Le principal résultat de cette décomposition en effet direct et indirect montre que l'innovation dans un département i n'est pas seulement influencée par la R&D de ce même département mais aussi par l'innovation des départements voisins.

Conclusion générale

Nous souhaitons dans cette thèse approfondir les recherches théoriques et empiriques sur la diffusion des connaissances en France métropolitaine. Notre ambition ne consistait pas à savoir si les connaissances diffusées sont des externalités de connaissance ou simplement des flux de connaissance, mais plutôt à étudier l'étendue géographique de ce processus de diffusion des connaissances en utilisant plusieurs approches empiriques. Ces approches par l'économétrie des données de panel, l'analyse exploratoire spatiale et l'économétrie spatiale se sont révélées très fructueuses. Nos résultats empiriques ont permis de mettre en avant les effets des connaissances localisées et distantes sur les activités innovantes en France, tout en soulignant l'importance des canaux de transmission de ces connaissances.

Cette thèse s'est déclinée en quatre chapitres. Le premier a présenté un survey théorique sur la question des frontières géographiques de la diffusion des connaissances ; les trois chapitres suivants, empiriques, ont répondu à notre problématique et ont développé nos hypothèses de recherche.

* * *

Dans **le premier chapitre**, nous sommes revenus sur le concept de "*connaissance*". Ce chapitre a permis de rappeler que la connaissance possède en partie les caractéristiques d'un bien public. Il en résulte des bénéfices collectifs qui enrichissent les connaissances des entreprises externes, d'où la notion d'externalités de connaissance. Le rôle de ces externalités dans les activités innovantes a pris une place importante dans la littérature économique (Cowan et al., 2000 ; Autant-Bernard et al, 2008).

Prenant appui sur cette littérature, nous avons mobilisé les approches de l'économie géographique, de l'économie de la connaissance et de la géographie de l'innovation afin de faire ressortir le caractère localisé de la connaissance. Ces approches ont largement souligné l'importance de la proximité géographique pour bénéficier des externalités de connaissance localisées (*Localized Knowledge Spillovers, LKS*) ou, plus généralement, des flux de connaissance. L'hypothèse qui sous-tend cette conclusion est qu'une partie

de la connaissance est tacite et que le face à face est une condition nécessaire pour son échange.

Toutefois, la nature tacite comme justification des LKS a été remise en cause dans la mesure où la transmission des connaissances tacites ne peut être que volontaire même en présence de contacts face à face (Breschi et Lissoni, 2003 ; Gallié, 2002).

Nous avons montré qu'au delà de sa nature tacite/codifiée, la connaissance présente un aspect collectif qui fait qu'elle peut être diffusée loin dans l'espace, d'où la notion d'externalités de connaissance distantes (*Distant Knowledge Spillovers, DKS*). En effet, les DKS jouent un rôle aussi important que les LKS dans la production d'innovation. Comme ces externalités ne sont pas bornées dans un espace géographique donné, elles peuvent se diffuser à une distance plus longue que les LKS.

Dans ce chapitre, nous avons également montré, à travers un panorama des mécanismes et des différents canaux de diffusion des connaissances, que si l'espace compte, c'est parce qu'il existe des réseaux d'innovation qui facilitent la collaboration et le transfert des connaissances.

Nous avons clos ce chapitre en précisant l'importance des pôles de compétitivité dans le transfert de connaissance. Les études empiriques sont loin d'être unanimes sur les méthodes de mesure des *spillovers* géographiques ainsi que sur l'ampleur des connaissances diffusées. Cela s'explique par le fait qu'elles étudient des canaux de diffusion des connaissances très variés et des formes d'agglomération des activités innovantes elles aussi très hétérogènes.

Le deuxième chapitre offre des prolongements empiriques au premier chapitre. Nous avons mobilisé deux principales approches empiriques pour étudier la relation entre espace et innovation par le biais des externalités de connaissance. La première approche a consisté à étudier les flux de connaissance appréhendés par les citations de brevets. Cette méthode a été par la suite critiquée (Almeida et Kogut, 1999). En effet, comme il existe un lien entre les citations de brevets et les brevets, toutes les critiques qui s'appliquent aux brevets s'appliquent aussi aux citations de brevets. La deuxième approche s'est appuyée sur une fonction de production de connaissance telle que proposée par Jaffe (1989). Cette dernière présente l'avantage, par rapport à la première, d'envisager l'interaction entre unités voisines ou proches.

Ensuite, partant des travaux de Bottazzi et Peri (2003) consacrés à l'estimation de la fonction de production de connaissance, nous avons proposé d'opérer une différenciation de la R&D en distinguant R&D interne, R&D externe privée et R&D externe publique. Nous avons estimé les effets de leurs *spillovers* géographiques sur le processus d'innovation.

Ces estimations ont été conduites en deux temps. Nous avons tout d'abord mené notre étude en nous basant sur des données en coupe. L'analyse a consisté à mesurer l'effet des *spillovers* des dépenses en R&D moyennes (mesurées entre 2002 et 2005) sur l'innovation du département considéré, cumulées entre 2005 et 2008. Ces *spillovers* ont été mesurés dans des classes de distances géographiques d'ampleur 200 kilomètres puis dans des classes de distances plus fines (d'ampleur 100 kilomètres) afin de mieux délimiter l'effet de ces *spillovers*. De plus, nous avons introduit dans le modèle plusieurs variables mesurant la qualité et la quantité des réseaux d'innovation. Deux résultats ont été tirés de ces estimations en coupe transversale : le premier montre que seule la R&D externe privée mesurée dans l'intervalle $]0 - 100km]$ a un effet significatif sur le dépôt de brevets. Ce résultat suggère que le transfert de connaissances codifiées ne nécessite pas une forte proximité géographique, contrairement aux connaissances tacites (mesurées par la R&D interne) qui nécessitent beaucoup plus de contacts face à face. Le deuxième résultat de ces estimations confirme l'importance des réseaux d'innovation dans la diffusion des connaissances.

Toutefois, nous avons précisé que ces estimations ne prennent pas en compte l'hétérogénéité entre observations. Afin d'apprécier la validité des estimateurs obtenus avec l'étude en coupe transversale, nous avons réalisé une autre analyse portant sur des données de panel. Cette dernière analyse est effectuée pour la période allant de 2002 à 2008. Les deux analyses font ressortir des résultats divergents en ce qui concerne l'effet et l'étendue des *spillovers* géographiques. Il ressort notamment que les effets de *spillovers* sur l'innovation sont à la fois négatifs et positifs selon le type de la R&D et l'intervalle considéré. L'approche par cercles concentriques de Bottazzi et Peri (2003) présente donc certaines limites qui se traduisent par une divergence de résultats entre l'estimation en coupe transversale et celle sur données de panel. Une estimation plus précise s'avère nécessaire pour mesurer l'ampleur des effets d'interdépendance spatiale entre les départements français.

Le troisième chapitre a permis de comparer les dynamiques spatiales de l'innovation de 2002 à celles de 2008. Nous avons mobilisé les techniques de mesure de l'autocorrélation spatiale permettant d'analyser les structures d'agglomération des activités innovantes en France. Il s'agit notamment des outils de l'analyse exploratoire des données spatiales, dite "*AEDS*". Par rapport aux indices classiques de concentration (notamment l'indice de Gini), l'AEDS a permis de tester la significativité de la concentration géographique en tenant compte explicitement de l'autocorrélation et de l'hétérogénéité spatiale de nos données. D'une part, l'AEDS a permis de mesurer l'agglomération globale à travers la statistique de Moran globale. D'autre part, elle s'est intéressée à l'analyse locale de l'autocorrélation à travers le diagramme de Moran,

les statistiques de LISA et les cartes de significativité de LISA. Cette dernière n'est qu'une combinaison des deux derniers indicateurs locaux. Le diagramme de Moran fournit l'état de chaque département français aux dates 2002 et 2008, et montre si chaque département correspond ou non au schéma global d'autocorrélation spatiale. Les statistiques de LISA ont pour objectif de montrer, si elles sont significatives, que l'innovation d'un département est liée à celle de ses voisins (association spatiale significative) ou au contraire, si elles ne le sont pas, que cette innovation est due à d'autres facteurs ne provenant pas des interactions existant entre les départements.

Compte tenu des données dont nous disposons, les résultats de l'analyse exploratoire confirment l'existence d'une dynamique spatiale positive et significative pour le cas des dépenses en R&D et des brevets. En effet, l'ensemble de l'analyse laisse apparaître la persistance de disparités relativement importantes au sein des départements français. Nous soulignons des discontinuités et des associations spatiales des brevets aussi bien positives que négatives. En 2002, 68,08% des associations (75,53% en 2008) appartiennent aux quadrants H-H (valeurs autocorrélées positivement et supérieures à la moyenne) et L-L (valeurs autocorrélées positivement et inférieures à la moyenne). Les départements atypiques sont représentés par les quadrants H-L (valeurs autocorrélées négativement et supérieures à la moyenne) et L-H (valeurs autocorrélées négativement et inférieures à la moyenne). En 2002, 31,92% (24,47% en 2008) des départements sont représentés par des associations L-H et H-L.

Cette disparité met en évidence les externalités de connaissance. Les externalités positives sont importantes dans la région Île-de-France. Comme elle occupe la première place dans les activités innovantes, cette région présente une dynamique spatiale positive indiquant une très forte dynamique d'agglomération du processus d'innovation. Cette constatation nous amène à préconiser une politique qui viserait l'accroissement des interactions et des complémentarités entre les périphéries. Dès lors, un retour sur la politique régionale et sur son rôle majeur quant à la réduction des inégalités au niveau des activités innovantes nous paraît pertinent.

Ces résultats, bien qu'étant exploratoires, ont été pris en compte dans nos estimations afin de déterminer l'ampleur de cette dépendance spatiale et l'étendue géographique des externalités de connaissance. Dans le prolongement de ce troisième chapitre, nous avons choisi, dans **le quatrième chapitre**, d'établir une modélisation basée sur des modèles spatiaux. Ce dernier chapitre a permis de mesurer de manière explicite les effets spatiaux. Nous avons proposé trois modèles spatiaux : un modèle autoregressif spatial à variable endogène décalée (SAR), un modèle autorégressif à variable endogène et à variables explicatives décalées spatialement (SDM), et enfin un modèle autoregressif sur les erreurs (SEM).

Dans la première partie de ce chapitre, nous avons présenté les résultats des estimations pour les trois modèles (SAR, SEM et SDM) sous trois spécifications différentes : spécification avec effets fixes individuels, spécification avec effets fixes temporels et spécification avec effets fixes individuels et temporels (mixte). Nous avons établi le test du multiplicateur de Lagrange appliqué sur nos données de panel spatial et ce qui nous a permis de choisir le modèle à variables décalées spatialement (SDM). En utilisant le critère d'information d'Akaike (AIC) le plus bas, nous avons retenu la spécification à effets mixtes du modèle (SDM).

Le coefficient d'autocorrélation spatiale ρ estimé est faible mais positif et significatif, ce qui indique la présence d'externalités positives émanant des brevets des départements voisins. Ces externalités influencent positivement l'innovation du département considéré. Il y a donc bien des *spillovers* de connaissance entre les départements français.

Par ailleurs, le modèle SDM a permis d'estimer les *spillovers* en provenance des variables explicatives, notamment les dépenses en R&D. Pour cela, nous avons proposé, dans la deuxième partie de ce même chapitre, de décomposer les effets des coefficients de ces variables en effets directs et effets indirects. Cela nous a permis de mesurer d'une manière plus explicite l'ampleur des *spillovers* liés aux variables explicatives. Les effets de *spillovers* sont représentés par les effets indirects et les effets feedbacks²⁹. L'ampleur de ces effets diminue rapidement avec la distance mais elle n'est pas négligeable pour les départements les plus proches. Les effets directs tiennent compte de l'effet de rétroaction. Ce dernier est le résultat de l'augmentation de l'innovation des départements voisins à un département i grâce à l'investissement de ce dernier dans les dépenses en R&D.

Nous avons démontré que la proximité géographique a des effets positifs sur la productivité de la R & D, que le contenu de la R&D interne en connaissances tacites lui donne un avantage sur la R&D externe, que la diversité industrielle des territoires a des effets positifs sur l'innovation, et que les pôles de compétitivité améliorent la productivité de l'innovation grâce à la richesse des canaux de transmission des connaissances dont ils disposent. Nous avons aussi proposé une décomposition spatiale de l'impact de la R&D en un effet direct et un effet indirect. Les effets *feedback* et indirect y représentent une mesure plus précise des externalités de connaissance.

* * *

29. Nous précisons que notre travail a été établi à un niveau agrégé (le département). Par conséquent, nous ne sommes pas en mesure d'évaluer les *spillovers* inter-entreprises.

Ces résultats de thèse pourraient être approfondis et améliorés sur plusieurs points. Tout d'abord, ils pourraient être enrichis par l'approfondissement de la question des externalités de connaissance négatives trouvées dans l'étude sur données de panel du chapitre 2. En général, la littérature considère que la proximité géographique a toujours un impact positif sur l'innovation. Les résultats de la deuxième partie du chapitre 2 nuancent cette idée. Il serait donc judicieux d'étudier la source de ces effets négatifs et de comprendre pourquoi la proximité géographique se transforme parfois en inconvénients et génère des freins à l'innovation.

Un deuxième prolongement envisagé pourrait consister à introduire dans le modèle spatial du chapitre 4 (le modèle SDM) des variables quantitatives mesurant d'une manière plus directe l'effet des canaux de transmission des connaissances sur l'innovation. En effet, la décomposition en effets directs et indirects ne s'applique que pour le cas d'un modèle à effets fixes, ce qui empêche l'estimation des coefficients des variables invariantes dans le temps. Ainsi, l'introduction de ces variables dans le modèle pourrait nous permettre d'obtenir des résultats plus appropriés dans la mesure où nous tiendrions compte à la fois de l'effet spatial et des variables réseaux.

Une dernière extension à ce premier travail de recherche serait d'établir notre étude à un niveau plus désagrégé notamment au niveau cantonal. Ceci permettrait de déterminer les dynamiques spatiales au niveau intra-départemental et donc d'étudier, d'une manière plus explicite, la disparité spatiale de l'innovation au sein de chaque département.

Bibliographie

- [1] Abdelmoula, M. et Bresson, G. (2007). Spatial and technological R&D spillovers in European regions : a dynamic count panel data model, *Annales d'Economie et de Statistique*, vol. 87-88, 167-194.
- [2] Acs, Z.J. et Audretsch, D.B. (1988). Innovation in large and small firms : an empirical analysis, *American Economic Review*, 78(4), 678-690.
- [3] Acs, Z. J. et Audretsch, D.B. (1990). *Innovation and Small Firms*. Cambridge, MA : MIT Press.
- [4] Acs, Z.J., Audretsch, D.B. et Feldman, M.P. (1992). Real Effects of Academic Research. *American Economic Review* 82(1), 363-367.
- [5] Acs, Z.J., Anselin , L. et Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge, *Research Policy* 31 (7) : 1069-1085.
- [6] Agarwal, R., Echambadi, R., Franco, A. M. et Sarkar, M. B. (2004). Knowledge transfer through inheritance : Spinout generation, development and survival, *Academy of Management Journal*, 47-4,501-522.
- [7] Agarwal, R., Chomsisengphet, S., Liu, C. et Souleles, N.(2006). Do Consumers Choose the Right Credit Contracts? Working Paper, Federal Reserve Bank of Chicago.
- [8] Agarwal, R., Audretsch, D. et Sarkar, M. (2010). Knowledge spillovers and strategic entrepreneurship, *Strategic Entrepreneurship Journal* 4 : 4, pp. 271-283.
- [9] Aghion, P. et Howitt, P. (1990). A model of growth through creative destruction, NBER Working Paper, n°3223.
- [10] Aiginger, K. (1999). Do industrial structures converge ? A survey on the empirical literature on specialisation and concentration of industries, WIFO-Working papers.

- [11] Almeida, P. et Kogut, B. (1997a). The exploration of technological diversity and the geographic localization of innovation, *Small Business Economics*, n° 9, p. 21-31.
- [12] Almeida, P. et Kogut, B. (1997b). The localization of ideas and the mobility of engineers in regional networks, Working Paper, juin, 45.
- [13] Almeida, P. et Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks, *Management Science*, 45(7), 905-917.
- [14] Amador, M. et Landier, A. (2003). Entrepreneurial pressure and innovation. Manuscript, MIT.
- [15] Amara, M. (2009). Concentration et agglomération industrielles dans la métropole de Tunis. *Revue d'Economie Régionale et Urbaine* 3, 599-624.
- [16] Amara, M. (2010). Inégalités spatiales et développement local en Tunisie. Thèse pour le doctorat en géographie et en gestion. Université de Paris I et Université de Tunis.
- [17] Amara, M., Kriaa, M. et Montacer, M. (2010). Centralité des activités dans la région urbaine de Tunis (1994-2004). *Revue d'Economie Régionale et Urbaine* 3, 473-509.
- [18] Amemiya, T. et MaCurdy, T. (1986). Instrumental-variable estimation of an error-components model, *Econometrica*, 54(4) : 869-880.
- [19] Amin, A. et Cohendet, P. (2004). *Architectures of knowledge : Firms, capabilities, and communities*, Oxford University Press, USA.
- [20] Amiti, M. (1998). New trade theory and industrial location in the EU : a survey of evidence, *Oxford Review Economic Policy*, vol (14), n°2.
- [21] Anselin, L. (1980). *Estimation Methods for Spatial Autoregressive Structures*, Cornell University, Regional Science Dissertation and Monograph Series #8, Ithaca, NY.
- [22] Anselin, L. (1988). *Spatial econometrics : methods and models*, Kluwer Academic Publishers, Dordrecht, The Netherlands, p284.
- [23] Anselin, L. et Rey, S. (1991). Properties of tests for spatial dependence in linear regression models, *Geographical Analysis*, 23 : 112-131.

-
- [24] Anselin, L. (1995). Local indicators of spatial association-LISA. *Geographical Analysis* 27 (2), 93-115.
- [25] Anselin, L. (1996). Interactive techniques and exploratory spatial data analysis. *Geographical Information Systems : Principles, Techniques, Management and Applications*.
- [26] Anselin, L., Varga, A. et Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations , *Journal of Urban Economics*, n 42, p. 422-448.
- [27] Anselin, L. et Bao, S. (1997). Exploratory Spatial Data Analysis Linking SpaceStat and ArcView. In M. Fischer and A. Getis (eds.), *Recent Developments in Spatial Analysis*. Berlin : Springer-Verlag, pp. 35-59.
- [28] Anselin, L. et Bera, A. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. *Handbook of applied economic statistics* 155.
- [29] Anselin, L. (1998a). Interactive Techniques and Exploratory Spatial Data Analysis, in *Geographical Information Systems : Principles, Techniques, Management and Applications*, Longley P.A., Goodchild M.F., Maguire D.J. et D.W. Wind (Eds.), Wiley, New York, 251-264.
- [30] Anselin, L. (1998b). Exploratory Spatial Data Analysis in a Geocomputational Environment, in *Geocomputation, a Primer*, Longley P.A., Brooks S.M., McDonnell R. et B. Macmillan (Eds.), New York, Wiley, 77-94.
- [31] Anselin, L. (1999). SpaceStat, a Software Program for the Analysis of Spatial Data, Version 1.90. Bio Medware, Ann Arbor.
- [32] Anselin, L., Cohen J., Cook, D., Gorr, W., et Tita, G. (2000). Spatial analyses of crime. In *Criminal justice : Volume 4, Measurement and analysis of crime and justice*. Washington, DC : National Institute of Justice.
- [33] Anselin, L. (2003). GeoDa™ 0.9 User's Guide, Revised, June 15, 2003.
- [34] Anselin, L. et Le Gallo, J. (2006). Interpolation of Air Quality Measures in Hedonic House Price Models : Spatial Aspects. *Spatial Economic Analysis* 1 (1), 31-52.
- [35] Anselin, L., Le Gallo, J. et Jayet, H. (2008). Spatial panel econometrics. *The Econometrics of Panel Data*, 625-660.

- [36] Anton, J. J. et Yao, D. (1995). Start-ups, Spin-offs, and Internal Projects, *Journal of Law, Economics and Organization* 11, 362–78.
- [37] Antonelli, C. (ed.)(1994). *The Economics of Standards, Special Issue, Information, Economics and Policy*, vol. 6.
- [38] Antonelli, C. (1999). *The microdynamics of technological change*, Londres, Routledge.
- [39] Antonelli, C. (2006). The governance of localized knowledge. An information economics approach to the economics of knowledge, *Industry and Innovation*, 13 (3), pp. 227–261.
- [40] Antonelli, C. et Calderini, M. (2008). The governance of knowledge compositeness and technological performance : the case of the automotive industry in Europe, *Economics of Innovation and New Technology*, 17(1), pp. 23 – 41.
- [41] Arbia, G. (2001a). Modelling the geography of economic activities on a continuous space. *Papers in Regional Science* 80 (4), 411-424.
- [42] Arbia, G. (2001b). The role of spatial effects in the empirical analysis of regional concentration. *Journal of Geographical Systems* 3 (3), 271-281.
- [43] Arora, A. et Gambardella, A. (1994). Evaluating Technological Information and Utilizing it, *Journal of Economic Behavior and Organization*, Vol.24, pp.91-114.
- [44] Arrow, K. (1962). The economic implications of learning by doing, *The review of economic studies* 29(3), 155-173.
- [45] Arundel, A. et Van de Paal G. (1995). *Innovation strategies of Europe's Largest Industrial Firms*, unpublished manuscript, MERIT.
- [46] Arundel, A. (2001). The relative effectiveness of patents and secrecy for appropriation , *Research Policy*, 30, 611-624.
- [47] Asheim, B., Coenen, L., Moodysson, J. et Vang, J. (2007). Constructing knowledge-based regional advantage : Implications for regional innovation policy, *International Journal of Entrepreneurship and Innovation Management* 7(2), 140-155.
- [48] Audretsch, D. B. et Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production, *American Economic Review*, 86(4) : 253-273.

- [49] Autant-Bernard, C. (2000). Géographie de l'innovation et externalités locales de connaissances : Une étude sur données françaises, Thèse pour le doctorat en sciences économiques, Université Jean Monnet St-Etienne.
- [50] Autant-Bernard, C. (2001a). The geography of knowledge spillovers and technological proximity. *Economics of Innovation and New Technology*, vol (10), 237-254.
- [51] Autant-Bernard, C. (2001b). Science and knowledge flows : Evidence from the French case. *Research Policy.*, 30(7), 1069-1078.
- [52] Autant-Bernard, C., Billand, P. et Massard, N. (2007). Social distance versus spatial distance in R&D cooperation. Empirical evidence from European collaboration choices in Micro and nanotechnologies. *Papers in Regional Science*, 86 (3), 495-519.
- [53] Autant-Bernard, C., LeSage, J.P et Parent, O. (2008). Specialization, Diversity and Geographical Diffusion of Knowledge, *Annales d'Économie et de Statistiques*, numéro spécial, *Spatial Econometrics, Innovative Networks and Growth*, 87-88.
- [54] Autant-Bernard, C. et LeSage, J.P. (2011). Quantifying Knowledge Spillovers Using Spatial Econometric Models, *Journal of Regional Science*, 51(3) : 471-496.
- [55] Ayadi, M. et Amara, M. (2009). Spatial Patterns And Geographic Determinants Of Welfare And Poverty In Tunisia, Working Papers 478, Economic Research Forum, revised Mar 2009.
- [56] Aydalot, P. (ed). (1984). Crise et espace, Paris, Economica.
- [57] Bagnasco, A. (1977). Tre Italie. La problematica territoriale dello sviluppo italiano, Il Mulino, Bologna.
- [58] Bailey, T.C. et Gatrell, A.C. (1995). *Interactive Spatial Data Analysis*. Addison Wesley Longman. Available at Ulrich's, Michigan Union, and Michigan Book and Supply.
- [59] Baldwin, W.L et Scott, J.T. (1987). Market Structure and technological Change. A volume in the *Economics of Technological Change* section, edited by F.M. Scherer, Harwood Academic Publishers, London, Paris, New York. Melbourne.
- [60] Balland, P.A. (2010). Proximity and the structure of knowledge networks : Evidence from the GNSS industry, Thèse de doctorat de Sciences Economiques, Université de Toulouse.

- [61] Balland, P.A., de Vaan, M. et Boschma, R. (2011). The Dynamics of Interfirm Networks along the Industry Life Cycle : The Case of the Global Video Games Industry 1987-2007, Papers in Evolutionary Economic Geography (PEEG) 1114, Utrecht University, Section of Economic Geography, revised Aug 2011.
- [62] Balland, P.A. (2012). Proximity and the Evolution of Collaboration Networks : Evidence from Research and Development Projects within the Global Navigation Satellite System (GNSS) Industry, *Regional Studies*, 46 (6) : 741-756.
- [63] Baltagi, B.H. (2005). Solutions to Problems Posed in Volume 20(1) and 20(2) : A Hausman Test Based on the Difference between Fixed Effects Two-Stage Least Squares and Error Components Two-Stage Least Squares Solution, *Econometric Theory*, Cambridge University Press, vol. 21(02), pages 483-484, April.
- [64] Baltagi, B.H, Bratberg, E. et Holmås, T.H. (2005). A panel data study of physicians labor supply : the case of Norway, *Health Economics*, John Wiley & Sons, Ltd., vol. 14(10), pages 1035-1045.
- [65] Baltagi, B.H. et Li, D. (2006). Prediction in the Panel Data Model with Spatial Correlation : the Case of Liquor, *Spatial Economic Analysis*, Taylor and Francis Journals, vol. 1(2), pages 175-185.
- [66] Bankman, J. et Gilson, R. (1999). Why start-ups ? *Stanford Law Review* ; 289-308.
- [67] Barlet , C., Duguet, E., Encaoua, D. et Pradel, J. (1998). The commercial success of innovations : an econometric analysis at the firm level in French manufacturing. *Annales d'Economie et de Statistique*, n 49-50, pp. 457-478.
- [68] Bathelt, H., Malmberg, A. et Maskell, P. (2004). Clusters and knowledge : local buzz, global pipelines and the process of knowledge creation, *Progress in Human Geography*, 28 (1) : 31-56.
- [69] Baumont, C., Ertur, C. et Le Gallo, J. (2003). Spatial Analysis of Employment And Population Density : The Case Of The Agglomeration Of Dijon, 1999 Urban/Regional 0310003, EconWPA.
- [70] Baumont, C., Ertur, C. et Le Gallo, J. (2001). A Spatial Econometric Analysis of Geographic Spillovers and Growth for European Regions, 1980-1995, LATEC - Document de travail - Economie (1991-2003) 2001-04, LATEC, Laboratoire d'Analyse et des Techniques Economiques, CNRS UMR 5118, Université de Bourgogne.

-
- [71] Beaudry, C. et Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate, *Research Policy*, 38, pp. 318-337.
- [72] Becattini, G. (1990). The Marshallian Industrial District as a Socio-Economic Notion. In Frank Pike, Giacomo Becattini, Werner Sengenberger, eds., *Industrial Districts and Inter-Firm Cooperation in Italy*. Geneva : IILS.
- [73] Becker, W. et Dietz, J. (2004). R&D cooperation and innovation activities of firms-evidence for the German manufacturing industry, *Research Policy*, Elsevier, vol. 33(2), pages 209-223, March.
- [74] Beer, C. et Riedl, A. (2012). Modelling Spatial Externalities in Panel Data : The Spatial Durbin model revisited. *Papers in Regional Science*, 91(2), 299–318.
- [75] Behrens, K. et Thisse, J.F. (2007). Regional economics : a new economic geography perspective, *Regional Science and Urban Economics* 37 (Regional Science and Urban Economics at 35. A Retrospective/Prospective Special Issue), July 2007, pp. 457-465.
- [76] Belderbos, R., Carree, M., Diederer, B., Lokshin, B. et Veugelers, R., (2004). Heterogeneity in R&D cooperation strategies. *International Journal of Industrial Organization*, in press.
- [77] Blommestein, H.J. (1983). Specification and Estimation of Spatial Econometric Models, *Regional Science and Urban Economics*, 13 : 251-270.
- [78] Bocquet, R. et Brossard, O., (2007). The variety of ICT adopters in the intra-firm diffusion process : Theoretical arguments and empirical evidence, *Structural Change and Economic Dynamics*, Elsevier, vol. 18(4), pages 409-437, December.
- [79] Bocquet, R. et Brossard, O., (2008). Adoption des TIC, proximité et diffusion localisée des connaissances, *Revue d'économie régionale et urbaine*, Armand Colin, vol. 0(3), 411-446.
- [80] Boots, B.N. et Tiefelsdorf, M. (2000). Global and Local Spatial Autocorrelation in Bounded Regular Tessellations. *Journal of Geographical Systems*, 2, 319-48.
- [81] Boschma, R.A. et Lambooy, J.G. (1999). The prospects of an adjustment policy based on collective learning in old industrial regions, *GeoJournal*, vol. 49, no. 4, pp. 391-399.

- [82] Boschma, R.A., Lambooy, J.G. et Schutjens, V.A.J.M. (2002). Embeddedness and innovation. In : M. Taylor & S. Leonard (eds.), *Embedded enterprise and social capital. International perspectives*, Aldershot : Ashgate, pp. 19-35.
- [83] Boschma, R.A. (2004). The competitiveness of regions from an evolutionary perspective, *Regional Studies*, vol. 38, no. 9, pp. 1001-1014.
- [84] Boschma, R. (2005). Proximity and innovation : a critical assessment. *Regional Studies* 39, 61-74.
- [85] Boschma, R. et Iammarino, S. (2009). Related Variety, Trade Linkages and Regional Growth, *Economic Geography*, vol. 85 (3), 289-311.
- [86] Boschma, R. et Frenken , K. (2010). The spatial evolution of innovation networks. A proximity perspective, in : R. Boschma and R. Martin (eds.) *The Handbook of Evolutionary Economic Geography*, Cheltenham : Edward Elgar, pp. 120-135.
- [87] Boschma, R. (2011). Regional branching and regional innovation policy, in : K. Kourtit, P. Nijkamp and R. R. Stough (eds.), *Drivers of innovation, entrepreneurship and regional dynamics*, Springer Verlag, Berlin/Heidelberg, pp. 359-368.
- [88] Boschma, R. et Frenken, K. (2011). The emerging empirics of evolutionary economic geography, *Journal of Economic Geography*, Oxford University Press, vol. 11(2), pages 295-307, March.
- [89] Boschma, R., Minondo, A. et Navarro, M. (2012). The emergence of new industries at the regional level in Spain. A proximity approach based on product-relatedness, *Papers in Evolutionary Economic Geography* 12.01, Utrecht University.
- [90] Bottazi, L. et Peri, G. (2003). Innovation and spillovers in regions : Evidence from european patent data., *European Economic Review* 47, 687-710.
- [91] Boufaden, N. et Plunket, A. (2008). Proximity and innovation : Do biotechnology firms located in the Paris Region benefit from localized technological externalities ? *Les Annales d'Economie et Statistique*, 87-88).
- [92] Boufaden, N. (2010). *Co-localisation, Externalités de Connaissance et Innovation des Entreprises : Cas de la Biotechnologie dans la région Ile-de-France.*, Thèse de doctorat de Sciences Economiques, Université Paris-Sud XI.

- [93] Bound, J., Cummins, C., Griliches, Z., Hall, B.H et Jaffe, A.(1984). Who Does R&D and Who Patents ?, NBER Chapters, in : R & D, Patents, and Productivity, pages 21-54 National Bureau of Economic Research, Inc.
- [94] Brachert, M., Kubis, A. et Titze, M. (2011). Related variety, unrelated variety and regional functions. Identifying sources of regional employment growth in Germany from 2003 to 2008. IWH Discussion Papers, Halle, Germany.
- [95] Brandsma, A.S. et Kelletaper, R.H. (1979). A biparametric approach to spatial autocorrelation, *Environment and Planning A*, 11, 51-58.
- [96] Breschi, S. et Lissoni, F. (2001). Knowledge Spillovers and Local Innovation Systems : A Critical Survey, *Industrial and Corporate Change*, Oxford University Press, vol. 10(4), pages 975-1005, December.
- [97] Breschi, S. et Lissoni, F. (2004). Knowledge networks from patent data : Methodological issues and research targets, KITeS Working Papers 150, KITeS, Centre for Knowledge, Internationalization and Technology Studies, Università Bocconi, Milano, Italy.
- [98] Breschi, S. et Lissoni, F. (2009). Mobility of skilled workers and co-invention networks : an anatomy of localized knowledge flows', *Journal of Economic Geography* 9(4), 439-468.
- [99] Brette, O. et Chappoz, Y. (2007). The French Competitiveness Clusters : Toward a New Public Policy for Innovation and Research ?, *Journal of Economic Issues*, XLI, 2, June, p. 391-398.
- [100] Breusch, T. et Pagan, A. (1980). The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics. *Review of Economic Studies*, 47, 239-253.
- [101] Brossard, O. et Vicente, J. (2010). Cognitive and relational distance in alliances networks : Evidence on the knowledge value chain in the European ICT sector, in Cooke P., De Laurentis C., MacNeill S. and Collinge C. (eds) *Platforms of Innovation : Dynamics of New Industrial Knowledge Flows*, Edward Elgar, forthcoming.
- [102] Brossard, O et Moussa, I. (2012). The trilogy of knowledge spillovers in French regions : a history of nature, channels and boundaries, *Papers in Evolutionary Economic Geography (PEEG)* 1207, Utrecht University, Section of Economic Geography, revised May 2012.

- [103] Buenstorf, G. (2009). Is commercialization good or bad for science? Individual-level evidence from the Max Planck Society, *Research Policy* 38(2), 281-292.
- [104] Cabrer-Borras, B. et Serrano-Domingo, G. (2007). Innovation and R&D spillover effects in Spanish regions : A spatial approach. *Research Policy*, 36(9) :1357-1371.
- [105] Cairncross, F. (1997). *The Death of Distance : How the Communications Revolution Will Change Our Lives*. London : Orion Publishing.
- [106] Callon, M. (1993). Variety and irreversibility in networks of technique conception and adoption. In D. Foray and C. Freeman (Eds.) *Technology and the Wealth of Nations : The Dynamics of Constructed Advantage*. London, Pinter Publishers : 232-268.
- [107] Callon, M. (1999). The role of lay people in the production and dissemination of scientific knowledge, *Science Technology and Society* 4(1), 81-94.
- [108] Camagni, R. et Crevoisier, O. (éds) (2000), *Les milieux urbains : innovation, système de production et ancrage*, Enquête GREMI V, Neuchâtel : EDES et GREMI, p360.
- [109] Caniëls, M.C.J. (2000). The geographic distribution of patents and value added accross european regions, Working paper, MERIT, Août, p.10.
- [110] Cantwell, J. et Vertova, G. (2004). Historical evolution of technological diversification. *Research Policy*, 33, 511-529.
- [111] Carrincazeaux, C., Lung, Y. et Rallet, A. (2001). Proximity and Localisation of corporate R&D Activities, *Research Policy*, 30(5), p777-789.
- [112] Carrincazeaux, C. (2011). Les dynamiques spatiales de l'innovation, à paraître dans *Industries, Innovation, Institutions*,
[http ://hal.archives-ouvertes.fr/hal-00394246/fr/](http://hal.archives-ouvertes.fr/hal-00394246/fr/).
- [113] Cassiman, B. et Ueda, M. (2006). Optimal Project Rejection and New Firm Start-ups, *Management Science*, INFORMS, vol. 52(2), pages 262-275, February.
- [114] Castells, M. (2002). *La Galaxie Internet*, Paris, Fayard, p 365.
- [115] Catin, M., Hanchane, S. et Kamal, A. (2007). Structure industrielle, externalités dynamiques et croissance locale au Maroc, *Région et Développement* 25, 45-63.

- [116] Chabault, D. (2011). L'apport de la théorie des parties prenantes 430 la gouvernance des pôles de compétitivité, *Vie et sciences économiques* (1), 39-57.
- [117] Chantelot, S., Peres, S. et Virol, S. (2010). The geography of French creative class : An exploratory spatial data analysis, *Cahiers du GREThA 2010-16*, Groupe de Recherche en Economie Théorique et Appliquée.
- [118] Cheynet, P. et Fadaïro, M. (1998). Les méthodes de mesure des externalités technologiques. Un aperçu des travaux économétriques, Working Paper for Programme CNRS Les enjeux économiques de l'innovation.
- [119] Christensen, C. (1993). The rigid disk drive industry : a history of commercial and technological turbulence. *Business History Review*, 67 : 531-588.
- [120] Cliff, A. et Ord, J. K. (1973). *Spatial Autocorrelation*, Londres, Pion.
- [121] Cliff, A. et Ord, J.K. (1981). *Spatial Processes : Models and Applications*. Pion Ltd.
- [122] Coe, D. T. et Helpman, E.(1995). International R&D Spillovers, *European Economic Review*, 39(5). 859-887.
- [123] Cohen, W. M. et Levinthal, D.A (1989). Innovation and learning : the two faces of R&D, *the Economics Journal*, Vol. 99, p.569-596.
- [124] Cohen, W. M. et Levinthal, D.A. (1990). Absorptive Capacity : A new perspective on learning and innovation. *Administrative Science Quarterly*, 35 : 128-152.
- [125] Cohen, W.M et Klepper, S. (1992). The Anatomy of Industry R&D Intensity Distributions, *American Economic Review*, American Economic Association, vol. 82(4), pages 773-99, September.
- [126] Cohen, W., Nelson, R.R. et Walsh, J.P. (2000). Protecting their intellectual assets : appropriability conditions and why US manufacturing firms patent (or not). WBER Working Paper Series, n 7552.
- [127] Cohendet, P. et Llerena, P. (1999). La conception de la firme comme processeur de connaissances, *Revue d'Economie Industrielle*, n 88, pp.211-236.
- [128] Cohendet, P. et Meyer-Krahmer, F. (2001). The Theoretical and Policy Implications of Knowledge Codification, *Research Policy*, 30 (9), pp. 1563-1591.
- [129] Cohendet, P., Creplet, F. et Dupouet, O. (2006). Gestion des connaissances, firmes et communautés de savoir, Paris, *Economica*, p205.

- [130] Cohendet, P. Farcot, M. et Pénin, J. (2006). Codification et industries culturelles : un espace de créativité et d'innovation, *Management International*, 10, pp. 65-91.
- [131] Coissard, S. (2007). Perspectives. La nouvelle économie géographique de Paul Krugman, *Revue d'Économie Régionale et Urbaine*(1), 111-125.
- [132] Colletis, G. (2010). Co-évolution des territoires et de la technologie : une perspective institutionnaliste, *Revue d'Économie Régionale et Urbaine*, p. 235-249.
- [133] Combes, P.P. (2000a). Marshall-Arrow-Romer Externalities and City Growth, CERAS, Working Paper, n 99-06.
- [134] Combes, P.P. (2000b). Economic Structure and Local Growth : France, 1984-1993, *Journal of Urban Economics*, vol. 47, pp. 329-355.
- [135] Combes, P.P. et Duranton, G. (2006). Labour pooling, labour poaching, and spatial clustering, *Regional Science and Urban Economics* 36(1), 1-28.
- [136] Conti S., Malecki E. et Oinas P.E. (1995). *The industrial Enterprise and its Environment : Spatial Perspective*, Aldershot : Avebury.
- [137] Cooke, P. (2001). From technopoles to regional innovation systems : the evolution of localised technology development policy', *Canadian Journal of Regional Science* 24(1).
- [138] Cooke, P. et Delaurentis, C. (2007). Trends and drivers of the knowledge economy in seven business sectors, internal research paper, EURODITE. Available at http://www.eurodite.bham.ac.uk/partners/WP_3.asp
- [139] Courlet, C. (1991). Les systèmes productifs localisés, de quoi parle-t-on ? in *Industries, territoires et politiques publiques*, L'Harmattan, Logiques Economiques, p.13-34.
- [140] Courlet, C. (1999). Territoires et développement, *Revue d'économie régionale et urbaine*, No 3.
- [141] Cowan, R., David, P. et Foray, D. (2000). The explicit Economics of Knowledge Codification and Tacitness , *Industrial and Corporate Change*, Vol. 9 (2), pp. 212-253.
- [142] Crépon, B. et Duguet, E. (1997). Estimating the Innovation Function from Patent Numbers : GMM on Count Panel Data, *Journal of Applied Econometrics*, 12 : 243-263.

-
- [143] Crevoisier, O. (2001). L'approche par les milieux innovateurs : état des lieux et perspectives, *Revue d'économie régionale et urbaine*, 1, pp. 153-166.
- [144] Crevoisier, O. et Jeannerat, H. (2009). Territorial knowledge dynamics : from the proximity paradigm to the multi-location paradigm, *European planning studies* (17)8, pp. 1223-1241.
- [145] Czarnitzki, D., Ebersberger, B. et Fier, A. (2007). The relationship between R&D collaboration, subsidies and R&D performance : Empirical evidence from Finland and Germany, *Journal of Applied Econometrics*, John Wiley & Sons, Ltd., vol. 22(7), pages 1347-1366.
- [146] d'Aubigny G. (2006). Dépendance spatiale et autocorrélation, in Droesbeke J.-J. Lejeune M. et Saporta G. *Analyse statistique des données spatiales*. Chapitre 2, pp. 17-45. Editions Technip, Paris.
- [147] Debreu, G. et Herstein, I.N. (1953). Nonnegative square matrices. *Econometrica*, 21 : 597-607.
- [148] Del Canto, J.G. et Gonzalez, I.S. (1999). A resource-based analysis of the factors determining a firm's R&D activities. *Research Policy* 28, 8, 891-905.
- [149] Dixit, A. et Norman, V. (1980). The Gains from Free Trade, *The Warwick Economics Research Paper Series (TWERPS) 173*, University of Warwick, Department of Economics.
- [150] Doloreux, D. (2002). What we should know about regional systems of innovation, *Technology in society* 24(3), 243-263.
- [151] Dormont, B. (1989). Introduction à l'Econométrie des Données de Panel, CNRS. Un autre classique, quoiqu'un peu ancien, sur le sujet, en français, mais sans aucun programme.
- [152] Dosi, G. (1988). The nature of the innovative process, in G. Dosi et al. (Ed), *Technical change and economic theory*. London : Printer Publishers.
- [153] Dubin, R. (1988). Spatial autocorrelation, *Review of Economics and Statistics* 70, 466-474.
- [154] Dubin, R.A. (1998). Spatial autocorrelation : a primer, *Journal of Housing Economics*, 7, 304-327.
- [155] Dubin, R., Pace, R.K. et Thibodeau, T.G. (1999). Spatial autoregression techniques for real estate data, *Journal of Real Estate Techniques*, 7, 79-95.

- [156] Duranton, G., Martin, P., Mayer, T. et Mayneris, F. (2008). Les pôles de compétitivité : que peut-on en attendre ?, éditions rue d'Ulm.
- [157] Easterly, W. et Levine, R. (2001). It is not Factor Accumulation : Stylized Facts and Growth Models. *World Bank Economic Review*, Vol. 15, n 2.
- [158] 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.
- [159] Elhorst, J. (2003). Specification and estimation of spatial panel data models. *International regional science review* 26 (3), 244.
- [160] Elhorst, J. (2010a). Applied spatial econometrics : raising the bar. *Spatial Economic Analysis* 5 (1), 9-28.
- [161] Elhorst, J. (2010b). Spatial panel data models. *Handbook of applied spatial analysis*, 377-407.
- [162] Englmann, F.C. et Walz, U. (1995). Industrial Centers and Regional Growth in the Presence of Local Inputs, *Journal of Regional Science*, 35, 3-27.
- [163] Ertur, C. et Koch, W. (2007). Growth, technological interdependence and spatial externalities : theory and evidence, *Journal of Applied Econometrics*, John Wiley & Sons, Ltd., vol. 22(6), pages 1033-1062.
- [164] Escribano, A., Fosfuri, A. et Tribó, J. (2009). Managing external knowledge flows : The moderating role of absorptive capacity, *Research Policy* 38(1), 96-105.
- [165] Favereau, O. (1998). Notes sur la théorie de l'information à laquelle pourrait conduire l'économie des conventions. In Petit (Ed.), *L'économie de l'information*. Paris : La Découverte.
- [166] Feldman, M. P, (1994a). *The Geography of Innovation* Kluwer Academic Publishers, Boston.
- [167] Feldman, M. P. et Audretsch, D. B., (1999). Innovation in cities : Science-based diversity, specialization and localized competition, *European Economic Review*, Elsevier, vol. 43(2), pages 409-429, February.
- [168] Feldman, M.P. (2003). The Locational Dynamics of the US Biotech Industry : Knowledge Externalities and the Anchor Hypothesis, *Industry and Innovation*, Vol. 10, n 3, pp. 311-328.

- [169] Feser, E. (1998). Old and New Theories of Industry Clusters, in Steiner, M. (ed) Clusters and Regional Specialisation, Pion Limited, London.
- [170] Fingleton, B., Lewney, R. et Pinelli, D. (1997). Regional Growth and Convergence, in COMMISSION EUROPÉENNE, The Single Market Review, subseries 6, Aggregate and Regional Impact, vol. 1, Kogan Page, London.
- [171] Fingleton, B. (1999). Estimates of Time to Economic Convergence : an Analysis of Regions of the European Union, *International Regional Science Review*, 22, p. 5-34.
- [172] Fingleton, B. (2000). Convergence : International Comparisons Based on a Simultaneous Equation Model with Regional Effects , *International Review of Applied Economics*, vol. 14, p. 285-305.
- [173] Fischer, M., Varga, A. (2003). Spatial knowledge spillovers and university research : Evidence from Austria, *The Annals of Regional Science*, 37(2) : 303-322.
- [174] Florax, R. et Folmer, H. (1992). Specification and Estimation of Spatial Linear Regression Models : Monte Carlo Evaluation of Pre-Test Estimators, *Regional Science and Urban Economics*, 22 : 405-432.
- [175] Foray, D. et Lundvall, B.A. (1996). The knowledge-based economy : from the economics of knowledge to the learning economy, in *Employment and growth knowledge-based economy*, OECD, Paris.
- [176] Foray, D. (2000). *L'économie de la connaissance*, Paris, Collection REPERES.
- [177] Fosfuri, A. et Ronde, T. (2004). High-tech clusters, technology spillovers, and trade secret laws', *International Journal of Industrial Organization* 22(1), 45-65.
- [178] Fotheringham, A. S., Brunson, C. et Charlton, M. (2002). *Geographically weighted regression : the analysis of spatially varying relationships*. Chichester, Willey.
- [179] Franco, A. et Filson, D. (2006). Spin-Outs : Knowledge Diffusion Through Employee Mobility, *Rand Journal of Economics* 37, 841-860.
- [180] Frenken, K., Van Oort, F.G. et Verburg, T. (2007). Related variety, unrelated variety and regional economic growth *Regional Studies*, 41 (5), 685-697.
- [181] Fritsch, M. et Lukas, R. (2001). Who cooperates on R&D?, *Research Policy*, Elsevier, vol. 30(2), pages 297-312, February.

- [182] Fujita, M. et Krugman, P. (2004). The new economic geography : Past, present and the future, *Papers in regional science* 83(1), 139-164.
- [183] Galliano, D. et Roux, P. (2006). Les inégalités spatiales dans l'usage des TIC. Le cas des firmes industrielles françaises, *revue économique*, volume 57, pp 1449-1475.
- [184] Gallié, E.P. et Guichard, R. (2002). The impact of ICT sophistication on geographically distant networks : the case of plasma physics as seen from France, *NPRN et Conference*, March 21-23 2002, Brighton, 29.
- [185] Gallié, E.P. (2003). Une grille d'analyse de l'usage des TIC dans les différentes étapes de la coopération technologique », *Sciences de la société*, n 59, mai ;
- [186] Gallié, E.P. (2004). Coopération, externalités de connaissance et géographie de l'innovation : Le cas du secteur des biotechnologies en France.
- [187] Geary, R. (1954). The contiguity ratio and statistical mapping. *The Incorporated Statistician*, 115-146.
- [188] Getis, A. et Ord, J.K., (1992). The analysis of spatial association by use of distance statistics, *Geographical Analysis*, 24, 189-206.
- [189] Getis, A. et Ord J.K. (1995). Local spatial autocorrelation statistics : distributional issues and an application, *Geographical Analysis*, 27, 286-306.
- [190] Gilly, J. (1987). Espaces productifs locaux, politique d'emploi des firmes et transformations du rapport salarial', *Industries et régions*, *Economica*, Paris, 45-53.
- [191] Giuliani, E. et Bell, M. (2005). The micro-determinants of meso-level learning and innovation : evidence from a Chilean wine cluster, *Research policy* 34(1), 47-68.
- [192] Giuri, P. et Mariani, M., (2007). Proximity of Inventors and Knowledge Flows, *Mimeo*.
- [193] Glaeser, E.L., Kallal, H.D., Scheinkman J.A. et Shleifer, A. (1992). Growth in Cities, *Journal of Political Economy*, 100 (6), 1126-1152.
- [194] Glaeser, E. (1998). Are cities dying? , *The Journal of Economic Perspectives* 12(2), 139-160.

- [195] Glasmeier, A. (1988). Factors governing the Development of High-tech Industries Agglomerations : A Tale of three Cities, *Regional Studies*, 22, 287-301.
- [196] Goto, A. et Nagata, A. (1996). Study of innovation process by a survey data, presented at the Workshop on Appropriability and Technological Opportunities at National Institute for Science and Technology Policy).
- [197] Griffith, D.A. (1992b). What is spatial autocorrelation? Reflections on the past 25 years of spatial statistics, *L'Espace Géographique*, 21, 265-280.
- [198] Griliches, Z.(1979). Issues in assessing the contribution of research and development to productivity growth, *Bell Journal of Economics* 10 (1), 92-116.
- [199] Griliches, Z. (1984). Introduction to R & D, Patents, and Productivity, NBER Chapters, in : R & D, Patents, and Productivity, pages 1-20 National Bureau of Economic Research, Inc.
- [200] Griliches, Z., Pakes, A. et Hall, B-H. (1987). The Value of Patents as Indicators of Inventive Activity, in *Economic policy and technological performance*. Eds. : Partha Dasgupta and Paul Stoneman. Cambridge : Cambridge U. Press, pp. 97-124.
- [201] Griliches, Z. (1990). Patent Statistics as Economic Indicators : A Survey. *Journal of Economic Literature* 28 (December) : 1661-707.
- [202] Griliches, Z. (1992).The Search for R&D Spillovers, *Scandinavian Journal of Economics*, vol. 94, pp. 29-47.
- [203] Griliches, Z. (1994).National Science Foundation Grants for Economics : Response, *Journal of Economic Perspectives*, American Economic Association, vol. 8(1), pages 203-05, Winter.
- [204] Griliches, Z. (1998), *Patent statistics as economic indicators : a survey'*, University of Chicago Press.
- [205] Gromb, D. et Scharfstein, D. (2002). *Entrepreneurship in Equilibrium*," NBER Working Papers 9001, National Bureau of Economic Research, Inc.
- [206] Grossman, G.M et Helpman, E. (1991). Quality Ladders and Product Cycles, *The Quarterly Journal of Economics*, MIT Press, vol. 106 (2), pages 557-86, May.
- [207] Grossman, G.M. et Helpman, E. (1992). *Protection for Sale*, Papers 162, Princeton, Woodrow Wilson School - Public and International Affairs.

- [208] Grossman, G.M et Helpman, E. (1995). International R&D Spillovers, *European Economic Review*, 39, pp. 859–887.
- [209] Guellec, D. (2009). *Economie de l'innovation*, Paris, Collection REPERES.
- [210] Guillain, R. et Huriot, J. (2000). Le rôle agglomératif des externalités d'informations, mythe ou réalité', Baumont C., Combes P.-P., Derycke P.-H. et Jayet H., *Economie géographique : approches théoriques et empiriques*, Paris, Economica.
- [211] Guillain, R., Le Gallo, J. et Boiteaux-Orain, C. (2004). The evolution of the spatial and sectoral patterns in Ile-De-France over 1978-1997, LEG - Document de travail - Economie 2004-02, LEG, Laboratoire d'Economie et de Gestion, CNRS UMR 5118, Université de Bourgogne.
- [212] Guillain, R., Le Gallo, J. et Boiteux-Orain, C. (2006). Changes in spatial and sectoral patterns of employment in Ile-de-France, 1978-1997. *Urban Studies* 43 (11), 2075{2098.
- [213] Guillain, R. et Le Gallo, J. (2007). Agglomeration and dispersion of economic activities in Paris and its surroundings : An exploratory spatial data analysis, LEG - Document de travail - Economie 2007-01, LEG, Laboratoire d'Economie et de Gestion, CNRS UMR 5118, Université de Bourgogne.
- [214] Hagedoorn, J. (2003). Sharing intellectual property rights-an exploratory study of joint patenting amongst companies, *Industrial and Corporate Change* 12(5), 1035-1050.
- [215] Haining, R. (1990). *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge : Cambridge University Press.
- [216] Hall, B. H, Jaffe, A. B et Trajtenberg, M., (2001).The NBER Patent Citations Data File : Lessons, Insights and Methodological Tools, CEPR Discussion Papers 3094, C.E.P.R. Discussion Papers.
- [217] Hall, B.H. et R.H. Ziedonis (2001). The Patent Paradox Revisited : An Empirical Study of Patenting in the US Semiconductor Industry, 1979-95, *RAND Journal of Economics*, 32(1) : 101-128.
- [218] Harhoff, D., Scherer, F.M. et Vopel, K. (1999). Citations, Family Size, Opposition and the Value of Patent Rights, Discussion Paper No. 97-27, Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim.

- [219] Hauser, M., Cushman, F., Young, L., Kang-Xingjin, R. et Mikhail, J. (2007). A dissociation between moral judgments and justifications, *Mind et Language* 22(1), 1-21.
- [220] Hausman, J. (1978). Specification Tests in Econometrics. *Econometrica*, 46, 1251-1271.
- [221] Hausman, J. A., Taylor, W. E. (1981). Panel data and unobservable individual effects, *Econometrica*, 49 :1377-1398.
- [222] Heijmans, R.D.H. et Magnus J.R. (1986a). Asymptotic normality of maximum-likelihood estimators obtained from normally distributed but dependent observations, *Econometric Theory*, 12, 374-412.
- [223] Heijmans, R.D.H. et Magnus J. R. (1986b). Consistent maximum-likelihood with dependent observations : the general (non-normal) case and the normal case, *Journal of Econometrics*, 32, 253-285.
- [224] Hellman, T. F. (2007). When do employees become entrepreneurs? *Management Science*, 53(6) :919-933.
- [225] Henderson, R.M. et Clark, K.B. (1990). Architectural innovation : The reconfiguration of existing product technologies and failure of established firms, *Administrative Science Quarterly*, 35 : 1, 9-30.
- [226] Henderson, V., Kuncoro, A. et Turner, M. (1995). Externalities and Industrial Development, *Journal of Political Economy*, vol. 103, n 5, pp. 1067-1090.
- [227] Houdebine, M. (1999). Concentration géographique des activités et spécialisation des départements français, *Economie et Statistiques*, 6-7, 189-204.
- [228] Huang, J. S. (1984). The autoregressive moving average model for spatial analysis. *Australian Journal of Statistics* 26, 169-178.
- [229] Iammarino, S., Piva, M., Vivarelli, M. et Von Tunzelmann, N.(2009). Technological Capabilities and Patterns of Cooperation of UK Firms : A Regional Investigation, IZA Discussion Papers 4129, Institute for the Study of Labor (IZA).
- [230] Jacobs, J. (1969). *The economy of cities*. Random House, New York.
- [231] Jaffe, A. (1986), Technological opportunity and spillovers of R&D : Evidence from firms, patents, profits, and market value, *American Economic Review* 75, 984.100.

- [232] Jaffe, A. (1989). Real Effects of Academic Research, *American Economic Review*, 79, 957-70.
- [233] Jaffe, A., Trajtenberg, M. et Henderson, R. (1992). Ivory Tower Versus Corporate Lab : An Empirical Study of Basic Research and Appropriability," NBER Working Papers 4146, National Bureau of Economic Research, Inc.
- [234] Jaffe, A., Trajtenberg, M., Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citation, *Quarterly Journal of Economics*, 108 (3), 577-598.
- [235] Jaffe, A. (1996). Trends and patterns in research and development expenditures in the United States. *Proceedings of the National Academy of Sciences of the United States of America* 93. 23. 12658-12663.
- [236] Jaffe, A. (1998). The importance of spillovers in the policy mission of the advanced technology program, *Journal of Technology Transfer* 23, 11.19.
- [237] Jaffe, A. , Fogarty, M. S. et Bruce Banks, A. (1998). Evidence from Patents and Patent Citations on the Impact of NASA and Other Federal Labs on Commercial Innovation, *Journal of Industrial Economics*, Wiley Blackwell, vol. 46(2), pages 183-205, June.
- [238] Jaffe, A.B. et Trajtenberg, M. (1999). International knowledge flows : Evidence from patent citations, *Economics of Innovation and New Technology* 8, 105-136.
- [239] Jaffe, A., Trajtenberg, M. et Fogarty, M. (2000). The meaning of patent citations : Report on the NBER/Case-Western Reserve survey of patentees, Technical report, National Bureau of Economic Research.
- [240] Jaffe, A., Henderson, R. et Trajtenberg, M. (2005). Patent Citations and the Geography of Spillovers : A Reassessment. *American Economic Review*.
- [241] Jayet, H. (1993). *Analyse spatiale quantitative : une introduction*. Economica, Paris.
- [242] Jayet, H. (2001). Économétrie et données spatiales : une introduction à la pratique, *Cahiers d'Economie et Sociologie Rurales*, n 58-59, pp. 105-129.
- [243] Karray, Z. (2001). *Coopération technologique des firmes et compétences pour innover. Une modélisation des choix appliquée à l'industrie française*, Thèse pour le doctorat en sciences économiques, Université de Toulouse.

- [244] Katz, M. L et Shapiro, C. (1985). Network Externalities, Competition, and Compatibility, *American Economic Review*, American Economic Association, vol. 75(3), p424-440, June.
- [245] Kelejian, H.H. et Purcha, I.R. (1998). A Generalized Spatial Two Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances, *Journal of Real Estate Finance and Economics*, 17(1) : 99-121.
- [246] Kelejian, H.H. et Purcha, I.R. (1999). A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model , *International Economic Review*, 40 : 509-533.
- [247] Kelejian, H.H, Tavlas, G. et Hondroyiannis, G. (2006). A Spatial Modeling Approach to Contagion Among Emerging Economies, *Open Economies Review*, Vol. 17, 423-441.
- [248] Kirat, T. et Lung, Y. (1999). Innovation and proximity territories as loci of collective learning processes, *European Urban and Regional Studies* 6(1), 27-38.
- [249] Klepper, S. et Sleeper S. (2005). Entry by Spinoffs, *Management Science*, INFORMS, vol. 51(8), pages 1291-1306, August.
- [250] Klepper, S. et Thompson, P. (2010). Disagreements and intra-industry spinoffs, *International Journal of Industrial Organization*, Elsevier, vol. 28(5), pages 526-538, September.
- [251] Kriaa, M. et Karray, Z. (2010). Innovation and R&D investment of tunisian firms, *the journal of business inquiry*.
- [252] Krugman, P. (1979). A Model of Innovation, Technology Transfer, and the World Distribution of Income, *Journal of Political Economy*, University of Chicago Press, vol. 87(2), pages 253-66, April.
- [253] Krugman, P. (1980). Scale Economies, Product Differentiation, and the Pattern of Trade, *American Economic Review*, American Economic Association, vol. 70(5), pages 950-59, December.
- [254] Krugman, P. (1981). Intra-industry Specialization and the Gains from Trade, *Journal of Political Economy*, University of Chicago Press, vol. 89(5), pages 959-73, October.

- [255] Krugman, P. (1991a). Increasing returns and economic geography, *Journal of Political Economy*, 99(3), June, 483-499.
- [256] Krugman, P. (1991b). *Geography and trade*, the MIT Press.
- [257] Krugman, P. (1991c). *Cities in space : three simple models*, Technical report, National Bureau of Economic Research.
- [258] Krugman, P. (1996). *Domestic Distortions and the Deindustrialization Hypothesis*, NBER Working Papers 5473, National Bureau of Economic Research, Inc.
- [259] Kubo, Y. (1995). Scale Economies, Regional Externalities, and the Possibility of Uneven Development, *Journal of Regional Science*, 35, 29-42.
- [260] Lafourcade, M. et Mion, G. (2007). Concentration, agglomeration and the size of plants. *Regional Science and Urban Economics* 37 (1), 46-68.
- [261] Lamoreaux , N.R. et Sokoloff, K.L. (1997). *Inventors, Firms and the Market for Technology : US Manufacturing in the Late Nineteenth and Early Twentieth Centuries*, National Bureau of Economic Research.
- [262] Laurent, T., Ruiz-Gazen, A. et Thomas-Agnan, C. (2012). *GeoXp : an R package for interactive exploratory spatial data analysis*, *Journal of Statistical Software*, vol. 47, n°2, 2012.
- [263] Le Gallo, J. (2002a). *Econométrie spatiale : l' autocorrelation spatiale dans les modèles de régression linéaire*. *Economie et prévision* 4 (155), 139-158.
- [264] Le Gallo, J. (2002b). *Disparités géographiques et convergence des régions européennes : une approche par l'économétrie spatiale*, thèse de doctorat en analyse et politiques économiques, Université de Bourgogne, Dijon.
- [265] LeSage, J.P et Pace, R.K. (2004). *Models for Spatially Dependent Missing Data*, *The Journal of Real Estate Finance and Economics*, Springer, vol. 29(2), pages 233-254, 09.
- [266] LeSage, J.P. et Pace, K.R. (2007). *A matrix exponential spatial specification*, *Journal of Econometrics*, Elsevier, vol. 140(1), pages 190-214, September.
- [267] LeSage, J. et Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman & Hall/CRC.

-
- [268] Licht, G. et Zoz, K. (1998). Patents and R&D : An Econometric Investigation Using Applications for German, European and US Patents by German Companies, *Annales d'Économie et de Statistique*, 49-50 : 329-360.
- [269] Lieberman, M. et Montgomery D. (1988). First mover advantages, *Strategic Management Journal*, 9, 41-58.
- [270] Longhi, C. et Quéré, M.(1991). Les microsystèmes productifs, in R. Arena et alii (Sous la dir.), *Traité d'économie industrielle*, Paris, Economica, pp. 356-372.
- [271] Longley, P. A. et Tobon, C. (2004). Spatial dependence and heterogeneity in patterns of hardship : an intra-urban analysis, *Annals of the Association of American Geographers*, 94, 503-519.
- [272] Lucas, E. (1988). On the Mechanics of Economic Development, *Journal of Monetary Economics*, 22, 342.
- [273] Lundvall, B.A. (1992). *National Systems of Innovation : Towards a Theory of Innovation and Interactive Learning*, London : Pinter.
- [274] Lung, Y. (1997). Organisation spatiale et coordination des activités d'innovation des entreprises, Rapport pour le Commissariat Général du Plan, octobre.
- [275] MacKenzie, D. et Spinardi, G. (1995). The tacit knowledge, weapons design and the uninvention of nuclear weapons, *American Journal of Sociology*, vol. 101, n 1.
- [276] Madeuf, B. et Lefebvre, G. (2005). Les groupes français dans la globalisation de la R&D, p. 193-207 dans Ministère de l'Economie et des Finances, *L'industrie en France et la mondialisation*.
- [277] Maftai, A. (2010). Les Externalites De Connaissances Localisees-Aspets Theoriques Et Empiriques, *CES Working Papers* 2, 42-59.
- [278] Maillat, D., Quévit , M., et Senn, L. (1993). Réseaux d'innovation et milieux innovateurs, dans D. Maillat, M. Quévit et L. Senn (dirs.). *Réseaux d'innovation et milieux innovateurs : un pari pour le développement régional*. Neuchâtel : EDES, 1-13.
- [279] Maillat, D. (1995). Territorial dynamic, innovative milieus and regional policy, *Entrepreneurship et Regional Development* 7(2), 157-165.

- [280] Maillat, D. (2003). Développement des systèmes territoriaux de production, compétitivité et innovation , Acte du colloque international de développement local : le développement local dans la perspective du développement humain, Campo Grande, Brésil 25-28 novembre 2003.
- [281] Maïresse, J. et Sassenou, M. (1991). Recherche-développement et productivité, un panorama des études économétriques sur données d'entreprises. STI Revue 8, 9-45.
- [282] Maïresse, J. et Mulkey, B.(2007). An Exploration of Local R&D Spillovers in France, Annales d'Economie et de Statistique, ENSAE, issue 87-88, pages 145-166.
- [283] Malmberg, A. et Maskell, P. (2006). Localized Learning Revisited, Growth and Change, Gatton College of Business and Economics, University of Kentucky, vol. 37(1), pages 1-18.
- [284] Mameli, F., Iammarino, S. et Boschma, R. (2012). Regional variety and employment growth in Italian labour market areas : services versus manufacturing industries, Papers in Evolutionary Economic Geography, Utrecht University, Section of Economic Geography.
- [285] Mangolte, P.A. (1997). La dynamique des connaissances tacites et articulées : une approche socio - cognitive, Économie appliquée, N°2, p. 105-134.
- [286] Manuel De Frascati. (2002). Méthode type proposée pour les enquêtes sur la recherche et le développement expérimental. OCDE.
- [287] Marshall, A. (1890). Principles of economics : an introductory volume, Londres, Macmillan.
- [288] Marshall, A. (1919). Industry and Trade, London, Macmillan and Co.
- [289] Marshall, A. (1920). Principles of Economics, 8th edition, New York : Macmillan and Company.
- [290] Martin, P. et Ottaviano, G.I.P. (1996). Growth and Location, CEPR Discussion Paper Series.
- [291] Martin, P. et Ottaviano, G.I.P. (1999). Growing Locations : Industry Location in a Model of Endogenous Growth, European Economic Review, 43, 281-302.

- [292] Massard, N. et Riou, S. (2002). L'impact des structures locales sur l'innovation en France : spécialisation ou diversité? , *Région et Développement*, vol. 16, n° 2, pp. 111-36.
- [293] Massard, N., Torre, A. et Crevoisier, O. (2004). Proximité géographique et innovation, dans Pecqueur, B. et Zimmermann, J.B., Paris, Hermès, pp155-180.
- [294] Maurseth, P.B. et Verspagen, B. (2002). Knowledge Spillovers In Europe. A Patent Citations Analysis, *Scandinavian Journal of Economics*, 104(4), pp. 531-545.
- [295] Mazzoleni, R. et Nelson, R. (1998). The benefits and costs of strong patent protection : a contribution to the current debate, *Research policy* 27(3), 273-284.
- [296] Millo, G. et Piras, G. (2012). splm : Spatial Panel Data Models in R. *Journal of Statistical Software*, 47(1), 1-38. URL.
- [297] Moen, J. (2000). Is Mobility of Technical Personnel a Source of R&D Spillovers? , NBER Working Paper, n° 7834.
- [298] Mohnen, P. (1990). R&D and productivity growth : A survey of the literature.
- [299] Monck, C.S.P., Porter, R.B., Quintas, P., Storey, D.J.et Wynarczyk, P., (1988). *Science Parks and the Growth of High Technology Firms*. Croom Helm, London.
- [300] Moran, P. (1948). The interpretation of statistical maps, *Journal of the Royal Statistical Society*, 10B, 243-251.
- [301] Moran, P. (1950a). Notes on continuous stochastic phenomena, *Biometrika*, 37, 17-23.
- [302] Moran, P. (1950b). A test for serial independence of residuals, *Biometrika*, 37, 178-181.
- [303] Moreno, R., Paci, R. et Usai, S. (2005). Geographical and sectoral clusters of innovation in Europe, *The Annals of Regional Science*, vol. 39, n° 4, pp. 715-739.
- [304] Morgan, K. (1997). The learning region : institutions, innovation and regional renewal, *Regional studies* 31(5), 491-503.
- [305] Motta, M. et Ronde, T. (2002). Trade secret laws, labor mobility, and innovations, Working Papers 08-2002, Copenhagen Business School, Department of Economics.

- [306] Moulaert, F. et Sekia, F. (2003). Territorial innovation models : a critical survey, *Regional studies* 37(3), 289-302.
- [307] Moussa, I. (2009). Une vision renouvelée du rôle des brevets et des citations de brevets dans le développement, *Les Cahiers de l'ATM*, n° 24, 51-68.
- [308] Mowery, D. C., Oxley, J. E. et Silverman, B. S. (1997). Technological overlap and interfirm cooperation : Implications for the resource-based view of the firm, *Research Policy*, forthcoming.
- [309] Murdoch, J. (1995). Actor-networks and the evolution of economic forms : combining description and explanation in theories of regulation, flexible specialization, and networks. *Environment and Planning A* 27, 731-757.
- [310] Mutl , J. et Pfaffermayr, M. (2011). The Hausman test in a Cliff and Ord panel model, *Econometrics Journal*, Royal Economic Society, vol. 14(1), pages 48-76, February.
- [311] Nadiri, I. (1993). Innovations and technological spillovers. NBER Working Paper No. 4423.
- [312] Negassi, S. (2004). R&D co-operation and innovation a microeconomic study on French firms. *Research Policy*, 33, pp. 365-382.
- [313] Nelson, R. et Winter, S. (1982). An Evolutionary theory of economic change, *Production Sets and Organizational Capabilities*, Belknap/Harvard, p. 59-65.
- [314] Nooteboom, B. (2000). Learning and innovation in organizations and economies. Oxford University Press, Oxford.
- [315] Nooteboom, B. (2001). From evolution to language and learning, in J. Foster and s.Metcalf (eds), *Frontiers of evolutionary economics : competition, self-organisation and innovation policy*, pp. 41-69.
- [316] O'brien, R. (1992). Global financial integration : the end of geography, Council on Foreign Relations Press.
- [317] OCDE, (2000). A new economy ? The changing role of innovation and information technology in growth, OCDE.
- [318] Odland, J.D. (1988). Spatial Autocorrelation. Sage Publications.
- [319] Oliveau, S. (2005). Periurbanisation in Tamil Nadu a quantitative approach. CSH occasional Paper, Delhi, n° 15, 90 p.

-
- [320] Ord, J. (1975). Estimation Methods for Models of Spatial Interaction , Journal of the American Statistical Association, 70 : 120-126.
- [321] Ordober, J. (1991). A patent system for both diffusion and exclusion', The journal of economic perspectives 5(1), 43-60.
- [322] OST (1998). The dynamics of S&T activities in the EU regions, TSER project 1093.
- [323] Owen-Smith, J. et Powell, W.W. (2003). Knowledge networks in the Boston biotechnology community, submitted for publication.
- [324] Pace, R. K. et Barry, R. (1997). Sparse spatial autoregressions. Statistics and Probability Letters 33, 291-297.
- [325] Pace, R. K. et Gilley, O.W. (1997). Using the spatial configuration of the data to improve estimation, Journal of Real Estate Finance and Economics, 14, 333-340.
- [326] Pace, R.K et LeSage, J.P. (2004). Chebyshev approximation of log-determinants of spatial weight matrices, Computational Statistics & Data Analysis, Elsevier, vol. 45(2), pages 179-196, March.
- [327] Paci, R. et Usai, S. (2000). Externalities, knowledge spillovers and the spatial distribution of innovation, CRENOS Working Paper, Mars, 28.
- [328] Paci, R. et Usai, S. (2009). Knowledge flows across European regions, The Annals of Regional Science, Springer, vol. 43(3), pages 669-690, September.
- [329] Pakes, A. et Nitzan, S. (1983). Optimum Contracts for Research Personnel, Research Employment, and the Establishment of 'Rival' Enterprises, Journal of Labor Economics, 1, 345-365.
- [330] Palivos, T. et Wang, P. (1993). Spatial Agglomeration and Endogenous Growth, Regional Science and Urban Economics, 26, 645-669.
- [331] Paul, M., Planes, B., Huiban, J.P., et Sevestre, P. (2000). Externalization or Spillovers? R&D Activity Components and the Firm Performances. Communication à la journée de travail UMR/ERUDITE, Changement technique, innovation et espace, 20 mars, Dijon.
- [332] Pénin, J. (2003). Patent policy : a need to focus both on appropriation and coordination failures, European Journal of Economic and Social Systems 16, 109-128.

- [333] Pénin, J. (2005). Patents versus ex post rewards : A new look, *Research Policy* 34(5), 641-656.
- [334] Peri, G. (2005). International Migrations : Some Comparisons and Lessons for the European Union, Working Papers 06-36, University of California at Davis, Department of Economics.
- [335] Perrat, J. (2000). Dynamiques des firmes et politiques de développement régional et local : éléments pour une lecture de la “géographie des interdépendances”. In Gilly J.P., Torre A. (dir.). *Dynamiques de proximité*. Ed. l’Harmattan, Collection Emploi, Industrie et Territoire, pp. 251-280.
- [336] Polanyi, M. (1958/1978). *Personal Knowledge*, Routledge and Kegan Paul, London.
- [337] Polanyi, M. (1966). *The tacit dimension*. Doubleday, New York.
- [338] Ponds, R., Oort, F., Frenken, K. (2010). Spillovers and university–industry collaboration : an extended knowledge production function approach, *Journal of Economic Geography*, 10 :231–255.
- [339] Porter, M.E. (1990). *The Competitive Advantage of Nations*, Free Press, New York.
- [340] Porter, M.E. (1998). *On Competition*, Boston : Harvard Business School.
- [341] Porter, M. (2000). Clusters and competition, in Gordon E. Clark, eds., *Oxford handbook of economic geography*. Oxford : Oxford University Press.
- [342] Powell, W.W. (1990). Neither Market Nor Hierarchy – Network Forms of Organization , *Research in Organizational Behaviour*, Vol. 12, 295-336.
- [343] Powell, W.W. (1996). inter-Organizational Collaboration in the Biotechnology Industry. *Journal of Institutional and Theoretical Economics* 120 (1) : 197–215.
- [344] Puga, D. (1998). European regional policy in light of recent location theories, working paper, 44.
- [345] Pyke, F. et Sengenberger, W. (1990). *Industrial Districts and Inter-firm Cooperation in Italy*. Ginevra, International Institute for Labour Studies.
- [346] Quatraro, F. (2010). Knowledge coherence, variety and economic growth : Manufacturing evidence from Italian regions, *Research Policy*, Elsevier, vol. 39(10), pages 1289-1302, December.

- [347] Rallet, A. et Torre, A. (2001). Proximité Géographique ou Proximité Organisationnelle? Une analyse spatiale des coopérations technologiques dans les réseaux localisés d'innovation, *Economie Appliquée*, LIV, 1, 147-171.
- [348] Rallet, A. et Torre, A. (2007). *Quelles proximités pour innover ?*, Vol. 41, Editions L'Harmattan.
- [349] Ramsey, J. B. (1969). Test for Specification error in Classical Linear Least Squares Regression Analysis, *Journal of the Royal Statistical Society, Series B*. 31, 350-371.
- [350] Rapport de la direction générale de la compétitivité, de l'industrie et des services ; DGCIS-EUROLIO : L'impact de la politique des pôles de compétitivité sur le développement des collaborations entre acteurs du processus d'innovation, October 2011.
- [351] Ratti, R., Bramanti, A. et Gordon, R., (eds) (1997). *The dynamics of innovative regions. The GREMI approach*. Aldershot, U.K. : Ashgate.
- [352] Rey, S.J. et Montouri, B.D. (1999). U.S. regional income convergence : a spatial econometric perspective, *Regional Studies*, 33, 145-156.
- [353] Richardson, G. B. (1972). The Organization of Industry , *The Economic Journal*, September, pp. 883-896.
- [354] Rocha, H.O. (2004). Entrepreneurship and Development : the role of clusters , *Small Business Economics*, Vol. 23, p. 363-400.
- [355] Romer, P. (1986). "Increasing Returns and Long-run Growth", *Journal of Political Economy*, 94, 1002-1037.
- [356] Romer, P. (1990). Endogenous Technological Change, *Journal of Political Economy*, 98, 71.S102.
- [357] Romer, P. (1993). Idea gaps and object gaps in economic development, *Journal of monetary economics* 32(3), 543-573.
- [358] Rosenberg, R. (1976). Benthic faunal dynamics during succession following pollution abatement in a Swedish estuary. *Oikos* 27 : 414-427.
- [359] Saxenian, A. (1994). *Regional Advantage. Culture and Competition in Silicon Valley and Route 128*, Cambridge : Harvard University Press.

- [360] Schankerman, M. et Pakes, A. (1986). Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period. *Economic Journal*, Vol. 96, pp. 1052–1077.
- [361] Scherer, F. M., Herzstein, S.E., Dreyfoos, A., Whitney, W., Bachmann, O., Pesek, P., Scott, C., Kelly, T. et Galvin, J. (1959). *Patents and the Corporation : A Report on Industrial Technology under Changing Public Policy*.
- [362] Scherer, F. (1965). Firm size, market structure, opportunity, and the output of patented inventions, *The American Economic Review* 55(5), 1097-1125.
- [363] Scherer, F. (1984). *Innovation and Growth : Schumpeterian Perspectives*. Cambridge, MA : MIT Press.
- [364] Scherer, F. (1991). Changing perspectives on the firm size problem, in Z.J. Acs and D.B. AUDRETSCH (eds.), *Innovation and Technological Change. An International Comparison*, University of Michigan Press, Ann Arbor.
- [365] Schmookler, J. (1966). *Invention and economic growth*, Harvard University Press Cambridge, MA.
- [366] Schumpeter, J. (1942). *Capitalisme, socialisme et démocratie*, Traduction française de Gaël Fain, Payot, Paris, 1961.
- [367] Schumpeter, J. (1966). *Ten Great Economists From Marx to Keynes*, London : George Allen and Unwin.
- [368] Schwartz, B. (1975). *Queuing and Waiting : Studies in the Social Organization of Access and Delay* (Univ of Chicago Press).
- [369] Scitovsky, T. (1954). Two concepts of External Economies, *Journal of Political Economy*, vol. LXII, 70-82.
- [370] Simon, H. (1982). *Models of bounded rationality : behavioural economics and business organization*, vol. 2.
- [371] Simonis, D. (2002). *The New Economic Geography : A Survey of the Literature*. Federal Planning Bureau Working Paper 16(02).
- [372] Sjöholm, F.(1996). International transfer of knowledge : the role of international trade and geographic proximity, *Weltwirtschaftliches Archiv*, 132, pp. 97-115.

- [373] Soulié, N. (2006). Technologies de l'Information et de la Communication et dynamiques des espaces ruraux : le cas de la région Midi-Pyrénées, *Revue d'Économie Régionale et Urbaine*(4), 531-550.
- [374] Steinmueller, W.E. (1999). Networked knowledge and knowledge based economies, *Telematica Institut* , Delft, February Studies, Vol. 1, N°2, pp. 23-42.
- [375] Stiglitz, J. (2000). Capital Market Liberalization, Economic Growth, and Instability, *World Development*, Elsevier, vol. 28(6), pages 1075-1086, June.
- [376] Stiglitz, J. (2002). Information and the Change in the Paradigm in Economics, *The American Economic Review* 92(3), 460-501.
- [377] Stople, M. (2001). Mobility of Research Workers and Knowledge Diffusion as Evidence in Patent Data. The case of Liquid Crystal Display Technology, Kiel Working paper n°1038.
- [378] Suire, R. et Vicente, J. (2002). La localisation des entreprises de la Net-économie entre contingences historiques et géographiques, *eEurope : la société européenne de l'information*, *Economica*.
- [379] Tappeiner, U., Tasser, E., Leitinger, G., Cernusca, A. et Tappeiner, G. (2008). Effects of historical and likely future scenarios of land use on above-and belowground vegetation carbon stocks of an Alpine valley, *Ecosystems* 11(8), 1383-1400.
- [380] Taylor, C. et Silberston, Z. (1973). *The Economic Impact of the Patent System : A Study of the British Patent System*, Cambridge, England : Cambridge University Press.
- [381] Théry, G. (1994). *Les autoroutes de l'information*, La Documentation Française, Paris.
- [382] Thompson, P. et Fox-Kean M., (2005). Patent Citations and the Geography of Knowledge Spillovers : A Reassessment. *American Economic Review*, 95(1), 450-460.
- [383] Thompson, P. (2006). Patent Citations and the Geography of Knowledge Spillovers : Evidence from Inventor- and Examiner-added Citations, *The Review of Economics and Statistics*, MIT Press, vol. 88(2), pages 383-388, May.
- [384] Tiuhman, M. et Anderson, P. (1986). Technological Discontinuities and Organizational Environments, *AJmniatvtivt Scwtct Qmarttr/j*, 31, 439-465.

- [385] Tobler, W. (1979). Cellular geography, in S. Gale et G. Olsson (Eds), *Philosophy in Geography*, 379-386, Reidel, Dordrecht.
- [386] Trajtenberg, M. (1990). Product Innovations, Price Indices and the (Mis) Measurement of Economic Performance, NBER Working Papers 3261, National Bureau of Economic Research, Inc.
- [387] Trigilia, C. (1986). Small-firm development and political subcultures in Italy', *European sociological review* 2(3), 161-175.
- [388] Varga, A. (1998). Local Academic Knowledge Spillovers and the Concentration of Economic Activity, Regional Research Institute, West Virginia University, Research Paper No. 9803, Morgantown.
- [389] Veltz, P. (2005). *Mondialisation, villes et territoires. L'économie d'archipel*, Paris, PUF, coll. *Economie en liberté* (2ème édition), p262.
- [390] Verspagen, B., Schoenmakers, W. (2000). The spatial dimension of knowledge spillovers in Europe : Evidence from patenting data, Paper presented at the AEA Conference on Intellectual Property Econometrics, Alicante, 19-20 April.
- [391] Veugelers, R. (1997). Internal R&D expenditures and External Technology Sourcing, *Research Policy*, 26, 3, 303-316.
- [392] Veugelers, R. et De Backer, K. (1999). Access to external knowledge : an empirical analysis of alliances as spillover channel, CERP Working Papers, n°21.
- [393] Virol, S. (2006). Les trois dimensions du processus d'intégration régionale en Europe : une approche par l'économétrie spatiale, *Cahier du GRES* n° 2006 – 16.
- [394] von Hippel, E. (1976). The dominant role of users in the scientific instrument innovation process. *Research Policy*, 5(3), 212-239. doi : 10.1016/0048-7333(76)90028-7.
- [395] von Hippel, E. (1988). *The Sources of Innovation* (New York : Oxford University Press).
- [396] Walz, U. (1996). Transport Costs, Intermediate Goods and Localized Growth, *Regional Science and Urban Economics*, 26, 671-695.
- [397] Whitaker, J. (1975). John Stuart Mill's Methodology, *The Journal of Political Economy*, 1033-1049.

- [398] Wiggins, S.N. (1995). Entrepreneurial Enterprises, Endogenous Ownership, and the Limits to Firm Size. *Economic Inquiry*, 33, (1) 54-69.
- [399] Zaninetti, JM.(2005). *Statistique spatiale. Méthodes et applications géomatiques*. Coll. Applications des SIG, Hermès Lavoisier, 320.
- [400] Zucker, L. et Darby, M. R. (1994). Intellectual Capital and the Birth of U.S. Biotechnology Enterprises. NBER Working Paper no 4653. NBER, Cambridge, Ma.
- [401] Zucker, L.G., Darby, M.R. et Armstrong, J. (1998). Geographically localized knowledge : Spillovers or markets ? *Economic Inquiry* 36, 65-86.

Annexes

Annexes

Annexe 1

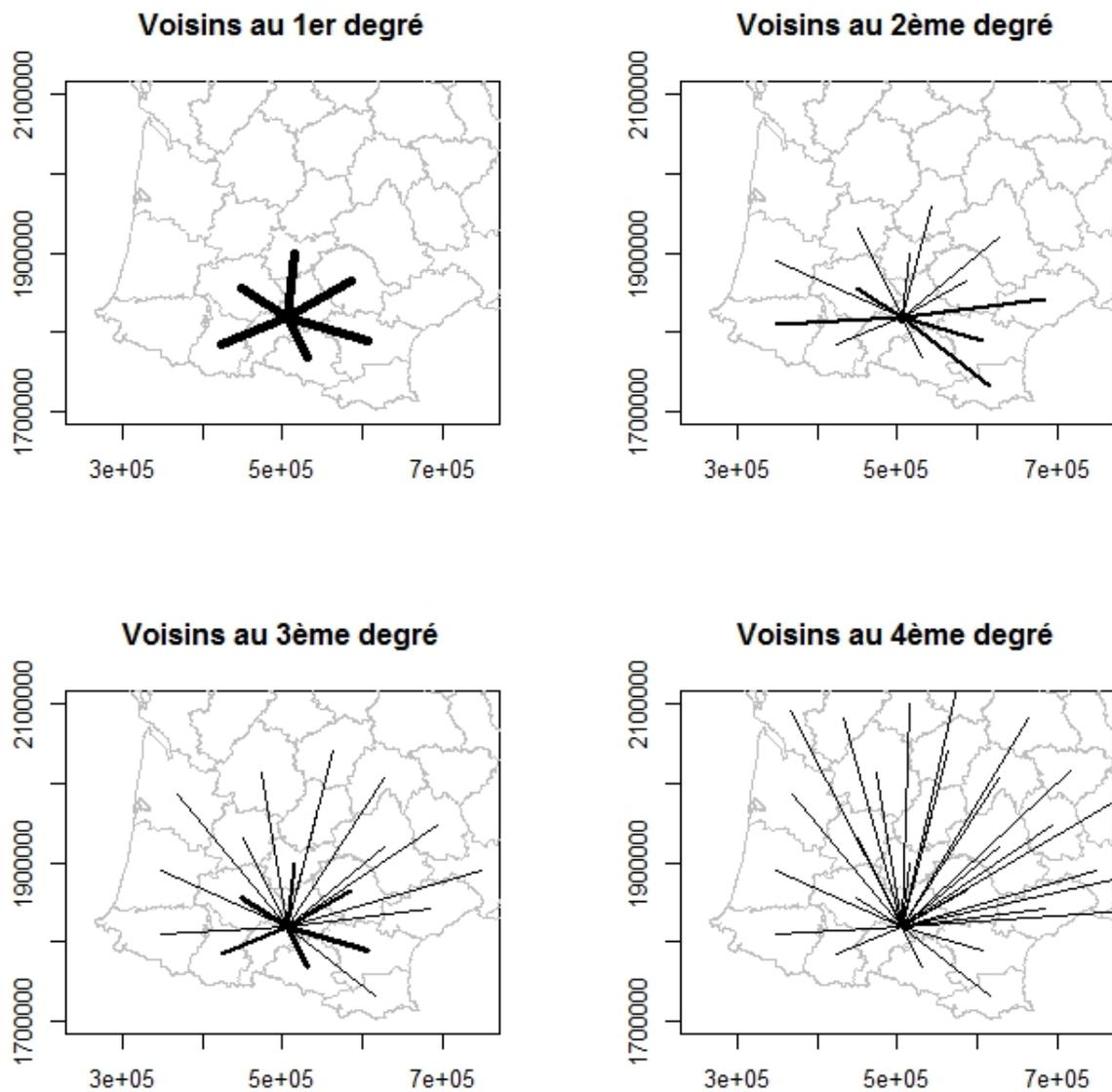


1. Carte des départements français métropolitains (Corse exclue)

(01) Ain, (02) Aisne, (03) Allier, (04) Alpes-de-Haute-Provence, (05) Hautes-Alpes
(06) Alpes-Maritimes, (07) Ardèche, (08) Ardennes, (09) Ariège, (10) Aube, (11)
Aude,(12) Aveyron, (13) Bouches-du-Rhône, (14) Calvados, (15) Cantal, (16) Charente,
(17) Charente-Maritime, (18) Cher, (19) Corrèze, (21) Côte-d'Or, (22) Côtes-d'Armor, (23)
Creuse,(24) Dordogne, (25) Doubs, (26) Drôme, (27) Eure, (28) Eure-et-Loir, (29) Finistère,
(30) Gard, (31) Haute-Garonne, (32) Gers, (33) Gironde, (34) Hérault, (35) Ile-et-Vilaine,
(36) Indre, (37) Indre-et-Loire, (38) Isère, (39) Jura, (40) Landes,(41) Loir-et-Cher, (42)
Loire, (43) Haute-Loire, (44) Loire-Atlantique, (45) Loiret,(46) Lot, (47) Lot-et-Garonne,
(48) Lozère, (49) Maine-et-Loire, (50) Manche, (51) Marne, (52) Haute-Marne, (53) Mayenne,
(54) Meurthe-et-Moselle, (55) Meuse, (56) Morbihan, (57) Moselle, (58) Nièvre, (59) Nord,
(60) Oise, (61) Orne, (62) Pas-de-Calais, (63) Puy-de-Dôme, (64) Pyrénées-Atlantiques, (65)
Hautes-Pyrénées, (66) Pyrénées-Orientales, (67) Bas-Rhin, (68) Haut-Rhin, (69) Rhône, (70)
Haute-Saône, (71) Saône-et-Loire, (72) Sarthe, (73) Savoie, (74) Haute-Savoie, (75) Paris, (76)
Seine-Maritime, (77) Seine-et-Marne, (78) Yvelines, (79) Deux-Sèvres, (80) Somme, (81) Tarn,

(82) Tarn-et-Garonne, (83) Var, (84) Vaucluse, (85) Vendée, (86) Vienne, (87) Haute-Vienne, (88) Vosges, (89) Yonne, (90) Territoire de Belfort, (91) Essonne, (92) Hauts-de-Seine, (93) Seine-Saint-Denis, (94) Val-de-Marne, (95) Val-d'Oise.

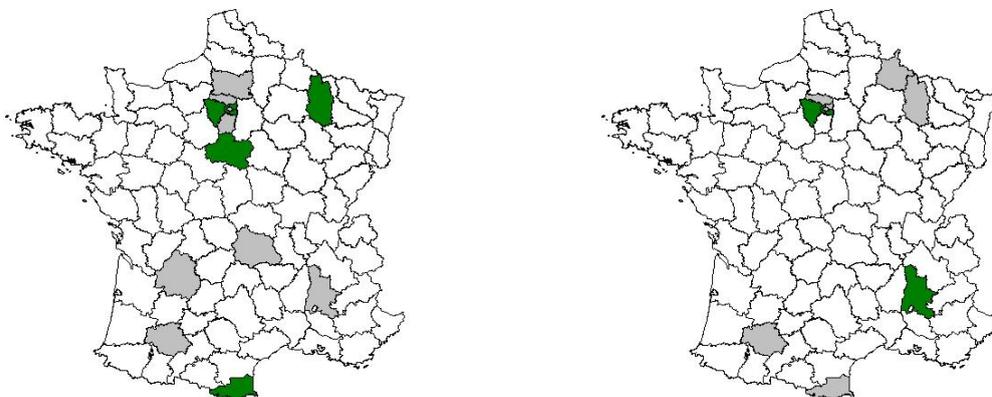
Annexe 2



2. Matrices de 1 à 4 voisins du département Haute-Garonne

Annexe 3

3.1 Cartes et tableaux de significativité de LISA pour le cas de *la R&D interne*



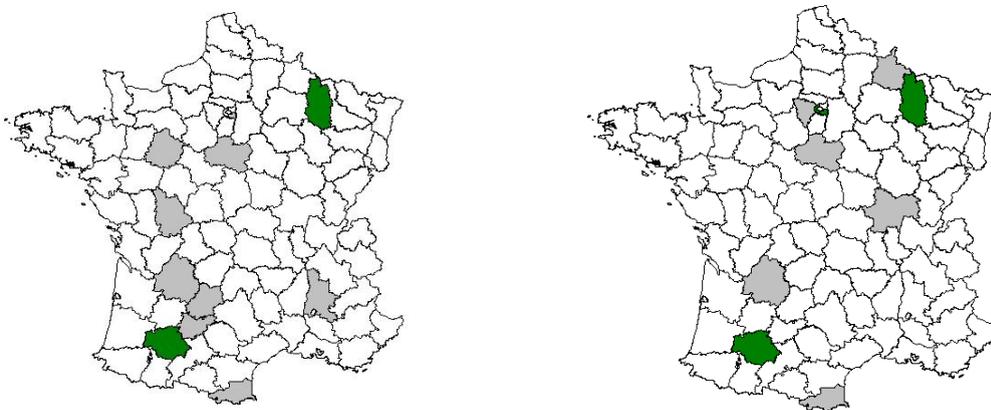
3.1.0 Carte de significativité de LISA de la R&D interne en 2002 et 2008

3.1.1 Statistiques de LISA de la R&D interne en 2002

Nom du département	code	Statistiques LISA	Clusters	p-value
Gers	32	0.179	L-L	0.01
Loiret	45	0.093	H-H	0.01
Meuse	55	0.189	L-L	0.01
Pyrénées-Orientales	66	0.196	L-L	0.01
Val-De-Marne	94	2.1028	H-H	0.01
Drôme	26	0.086	L-L	0.02
Oise	60	-0.000	L-H	0.02
Hauts-De-Seine	92	5.646	H-H	0.02
Val-D'oise	95	0.136	H-H	0.02
Seine-Saint-Denis	93	0.662	H-H	0.03
Yvelines	78	6.293	H-H	0.04
Essonne	91	1.905	H-H	0.04
Dordogne	24	0.146	L-L	0.05
Puy-De-Dome	63	-0.230	H-L	0.05

3.1.2 Statistiques de LISA de la R&D interne en 2008

Nom du département	code	Statistiques LISA	Clusters	p-value
Val-De-Marne	94	3.314	H-H	0.002
Yvelines	78	6.072	H-H	0.008
Hauts-De-Seine	92	6.369	H-H	0.008
Pyrénées-Orientales	66	0.343	L-L	0.012
Drôme	26	0.213	L-L	0.014
Loiret	45	0.260	H-H	0.014
Meuse	55	0.311	L-L	0.016
Val-D'Oise	95	0.378	H-H	0.016
Ardennes	08	0.308	L-L	0.022
Seine-Saint-Denis	93	1.303	H-H	0.034
Gers	32	0.275	L-L	0.044

3.2 Cartes et tableaux de significativité de LISA pour le cas de *la R&D externe publique*

3.2.0 Carte de significativité de LISA de la R&D externe publique en 2002 et 2008

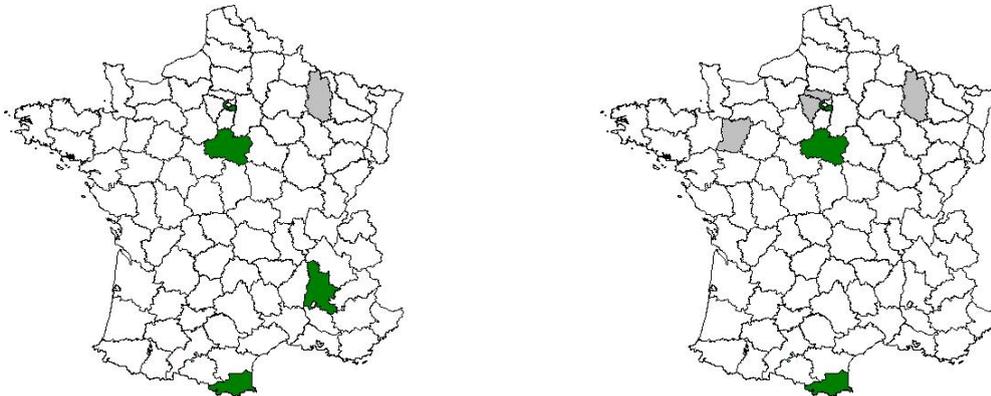
3.2.1 Statistiques de LISA de la R&D externe publique en 2002

Nom du département	code	Statistiques LISA	Clusters	p-value
Gers	32	0.118	L-L	0.002
Meuse	55	0.119	L-L	0.002
Val-De-Marne	94	0.0665	H-H	0.004
Pyrénées-Orientales	66	0.118	L-L	0.016
Loiret	45	-0.281	L-H	0.02
Lot	46	0.113	L-L	0.032
Dordogne	24	0.109	L-L	0.034
Vienne	86	-0.285	L-H	0.034
Drôme	26	0.033	L-L	0.042
Tarn-Et-Garonne	82	0.106	L-L	0.048
Sarthe	72	-0.182	L-H	0.038

3.2.2 Statistiques de LISA de la R&D externe publique en 2008

Nom du département	code	Statistiques LISA	Clusters	p-value
Meuse	55	0.189	L-L	0.001
Val-De-Marne	94	-0.150	L-H	0.001
Gers	32	0.182	L-L	0.003
Hauts-De-Seine	92	6.891	H-H	0.004
Loiret	45	-0.287	L-H	0.012
Yvelines	78	4.203	H-H	0.028
Ardennes	8	0.181	L-L	0.032
Pyrénées-Orientales	66	0.190	L-L	0.036
Saône-Et-Loire	71	-0.143	H-L	0.042
Dordogne	24	0.164	L-L	0.046

3.3 Cartes et tableaux de significativité de LISA pour le cas de *la R&D externe privée*



3.3.0 Carte de significativité de LISA de la R&D externe privée en 2002 et 2008

3.3.1 Statistiques de LISA de la R&D externe privée en 2002

Nom du département	code	Statistiques LISA	Clusters	p-value
Val-De-Marne	94	0.145	H-H	0.002
Loiret	45	-0.377	L-H	0.004
Drôme	26	0.062	L-L	0,01
Pyrénées-Orientales	66	0.102	L-L	0.01
Hauts-De-Seine	92	6.069	H-H	0.01
Yvelines	78	6.515	H-H	0.02
Meuse	55	0.098	L-L	0.026

3.3.2 Statistiques de LISA de la R&D externe privée en 2008

Nom du département	code	Statistiques LISA	Clusters	p-value
Hauts-de-Seine	92	6.411	H-H	0.01
Val-De-Marne	94	3.162	H-H	0.01
Loiret	45	0.259	H-H	0.02
Yvelines	78	6.632	H-H	0.02
Val-D'oise	95	-0.001	L-H	0.02
Meuse	55	0.258	L-L	0.03
Mayenne	53	0.185	L-L	0.05

Annexe 4

4. Spécification spatiale en panel à effets aléatoire

Modèle	RE-SEM		RE-SAR		RE-SDM	
$\ln R\&D_{int,it} - 1$	0.183**** (0.025)	0.185**** (0.025)	0.089**** (0.018)	0.099**** (0.019)	0.092**** (0.018)	0.098**** (0.018)
$\ln R\&D_{extpub,it} - 1$	0.025** (0.012)	0.022* (0.012)	0.023** (0.010)	0.022** (0.011)	0.024** (0.010)	0.023** (0.010)
$\ln R\&D_{extpriv,it} - 1$	0.037*** (0.013)	0.037*** (0.013)	0.018 (0.011)	0.022** (0.011)	0.015 (0.011)	0.020* (0.011)
$EG_{index,it}$	0.119**** (0.025)	0.121*** (0.025)	0.131**** (0.024)	0.134**** (0.024)	0.134**** (0.024)	0.135**** (0.024)
$EG_{0200,it}$	0.015 (0.031)	0.012 (0.031)	0.025 (0.022)	0.025 (0.023)		
$intercopubliscore_i$	0.162**** (0.047)		0.140**** (0.040)		0.139**** (0.040)	
$poleM_i$		0.697**** (0.088)		0.590**** (0.078)		0.590**** (0.077)
$largecity_i$	0.671* (0.395)	0.308 (0.308)	1.078*** (0.337)	0.750*** (0.272)	1.077*** (0.334)	0.764*** (0.270)
$W \times \ln(RD_{int,i,t-1})$					-0.021 (0.021)	-0.016 (0.022)
$W \times \ln(RD_{extpub,i,t-1})$					0.012 (0.014)	0.016 (0.015)
$W \times \ln(RD_{extpriv,i,t-1})$					0.025 (0.020)	0.025 (0.020)
$W \times EG_{index,it}$					0.015* (0.008)	0.015* (0.008)
$W \times intercopubliscore_i$					-0.004* (0.002)	
$W \times poleM_i$						-0.013* (0.007)
$W \times largecity_i$					0.006 (0.020)	-0.002 (0.019)
ρ			0.630*** (0.080)	0.549*** (0.088)	0.633*** (0.08)	0.547*** (0.088)
λ	0.229** (0.093)	0.234** (0.093)				
AIC	623.832	584.720	598.294	564.537	599.941	564.856

*p<0.1, **p<0.05, ***p<0.01, ****p<0.001

Annexe 5

Le problème du jacobien de l'estimation par le maximum de vraisemblance

Pour le modèle SEM (equation 4), le point de départ est l'hypothèse de normalité des termes d'erreur :

$\mu \sim Nid(0, \sigma^2 I)$. Puisque dans (5e6),

$\mu = (I - \lambda W_2)(y - \rho W_1 y - X\beta)$, le jacobien de la transformation est :

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

La fonction de log-vraisemblance de y s'écrit donc :

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

De cette expression, on déduit l'espace des paramètres de ρ et λ .

Soient ω_{max} la valeur propre positive maximale et ω_{min} la valeur propre négative la plus grande en valeur absolue de W .

Alors on démontre que $|I - \rho W_1|$ et $|I - \lambda W_2|$ sont positifs si l'inégalité $\frac{1}{\omega_{min}} < \psi < \frac{1}{\omega_{min}}$ (avec $\psi = \rho$ ou λ) est respectée.

Pour une matrice standardisée en ligne, $\omega_{max} = 1$ et $\omega_{min} > -1$.

Le problème principal de l'estimation par le maximum de vraisemblance est la présence du jacobien puisque la maximisation de la fonction de log-vraisemblance nécessite son évaluation pour chaque nouvelle valeur de ρ ou de λ . Il s'agit d'une opération lourde, même pour de petits échantillons, dans la mesure où il faut calculer le déterminant d'une matrice carré de dimension N .

Ord (1975) propose une solution qui consiste à exploiter la décomposition du jacobien en termes des N valeurs propres de la matrice de poids W .

$$|I - \rho W| = \prod_{i=1}^N (I - \rho \omega_i) \Rightarrow \ln |I - \rho W| = \sum_{i=1}^N \ln(1 - \rho \omega_i)$$

Ainsi, au lieu de calculer le déterminant de $I - \rho W$ à chaque étape, il n'est besoin de calculer qu'une fois pour toutes les valeurs propres de W si l'on utilise cette formulation simplifiée. La fonction de vraisemblance globale se réduit alors à la somme des vraisemblances individuelles. Certains auteurs fournissent des codes pour quelques logiciels économétriques basés sur cette propriété (Anselin et Hudak, 1992 ; Griffith, 1993 ; Li, 1996).

Le système d'équations résultant des conditions du premier ordre n'admet pas dans ce cas de solution analytique.

En revanche, les systèmes correspondant aux modèles (3 et 4) admettent des solutions issues des conditions du premier ordre permettant de construire une fonction de log-vraisemblance concentrée. On pourra se référer à Anselin (1988a) et Le Gallo (2000) pour plus de détails sur les formes prises par les estimateurs dans ces deux cas particuliers.

Table des figures

1.	Architecture de la thèse	16
1.1.	Les retombées privées et sociales : cas des externalités de connaissance	29
1.2.	Production et mesure de l'innovation	43
2.1.	Comparaison entre brevets de contrôle (B.Cr), brevets originels (B.O) et citations de brevets (B.Cit)	75
2.2.	Mesure des externalités dans des intervalles de 200 kilomètres	98
2.3.	Répartition du nombre moyen de brevets mesuré entre 2006 et 2008	107
2.4.	Répartition des dépenses moyennes en R&D interne mesurées entre 2002 et 2005	108
2.5.	Répartition des dépenses moyennes en R&D externe privée mesurées entre 2002 et 2005	109
2.6.	Répartition des dépenses moyennes en R&D externe publique mesurées entre 2002 et 2005	109
2.7.	Évolution du nombre des brevets entre 2002 et 2008	121
2.8.	Évolution des dépenses internes en R&D entre 2002 et 2008	122
2.9.	Évolution des dépenses externes publiques en R&D entre 2002 et 2008	122
2.10.	Évolution des dépenses externes privées en R&D entre 2002 et 2008	123
3.1.	Concentration des brevets et de la R&D interne et externe en 2002	138
3.2.	Concentration des brevets et de la R&D interne et externe en 2008	139
3.3.	Agglomération ou concentration ?	141
3.4.	Différentes formes de contiguïté	145
3.5.	Matrices de poids utilisées	148
3.6.	Les outils de l'analyse exploratoire de données spatiales (AEDES)	151
3.7.	Le diagramme de Moran	157
3.8.	Diagramme de Moran des brevets en 2002	162

TABLE DES FIGURES

3.9. Diagramme de Moran des brevets en 2008	162
3.10. Cartes de significativité de Moran des brevets en 2002	163
3.11. Cartes de significativité de Moran des brevets en 2008	164
3.12. Cartes de significativité de Moran des dépenses en R&D interne en 2002 et 2008	167
3.13. Cartes de significativité de Moran des dépenses en R&D externe publique en 2002 et 2008	169
3.14. Cartes de significativité de Moran des dépenses en R&D externe privée en 2002 et 2008	169
4.1. Répartition des effets directs et indirects de la R&D interne entre départements de la région Midi-Pyrénées	201

Liste des tableaux

1.1. Classification des connaissances en fonction de leur degré de codification et de manifestation	37
2.1. Statistiques descriptives pour les 94 départements français	113
2.2. Résultats d'estimation des modèles sans externalités géographiques	115
2.3. Résultats d'estimation des modèles avec externalités géographiques mesurées dans des intervalles de 200 kilomètres	116
2.4. Résultats d'estimation des modèles avec externalités géographiques mesurées dans des intervalles de 100 kilomètres	118
2.5. Statistiques descriptives des variables utilisées	125
2.6. Résultats d'estimation de la FPI sur des données de panel sans <i>spillovers</i> extra-départementaux	128
2.7. Résultats d'estimation de la FPI sur des données de panel avec <i>spillovers</i> extra-départementaux.	130
3.1. Indice de Gini	139
3.2. Répartition du nombre de voisins par département	149
3.3. Nombre total et moyen de voisins par département	150
3.4. Indice de Moran	155
3.5. Différences entre méthodes globales et méthodes locales de l'autocorrélation spatiale	156
3.6. Statistiques de LISA des brevets en 2002	164
3.7. Statistiques de LISA des brevets en 2008	165
4.1. Estimation du modèle SAR avec effets fixes	186
4.2. Estimation du modèle SEM avec effets fixes	187
4.3. Estimation du modèle SDM avec effets fixes	188
4.4. Test du Multiplicateur de Lagrange	189

LISTE DES TABLEAUX

4.5. Estimation des effets directs, indirects et totaux du modèle SDM (Effets mixtes du tableau 4.3)	197
4.6. $\ln(R\&Dint_{it-1})$: Répartition de l'effet direct, indirect et total par matrice de contiguïté d'ordre 0 à 4	199
4.7. $\ln(R\&Dextpub_{it-1})$: Répartition de l'effet direct, indirect et total par matrice de contiguïté d'ordre 0 à 4	199
4.8. $\ln(R\&Dextpiv_{it-1})$: Répartition de l'effet direct, indirect et total par matrice de contiguïté d'ordre 0 à 4	200
4.9. $EGindex_{it}$: Répartition de l'effet direct, indirect et total par matrice de contiguïté d'ordre 0 à 4	200

Table des matières

Introduction générale	11
1. Un éclairage théorique sur la question des frontières géographiques de la diffusion des connaissances	19
1.1. Introduction	19
1.2. De la nature de la connaissance aux Knowledge Spillovers (KS)	22
1.2.1. Propriétés économiques de la connaissance et externalités	23
1.2.1.1. Propriétés économiques de la connaissance	23
1.2.1.2. Définitions et source des externalités de connaissance	25
1.2.2. La nature tacite de la connaissance comme justification des Localized Knowledge Spillovers (LKS)	30
1.2.3. Remise en cause de la dimension tacite des connaissances comme justification des LKS	31
1.2.3.1. Le débat sur la notion de LKS	31
1.2.3.2. Transmission des connaissances tacites et contact face à face	33
1.2.3.3. La codification des connaissances tacites peut atténuer les LKS	34
1.2.4. La nature variée de la connaissance comme justification des Distant Knowledge Spillovers (DKS)	35
1.2.4.1. Distinction entre connaissances articulées et connaissances non articulées	36
1.2.4.2. Nature collective et composite de la connaissance	38
1.3. L'innovation : une activité concentrée géographiquement, mais dans des frontières mouvantes	39
1.3.1. Définition et mécanismes de l'innovation	40
1.3.1.1. Aspect interactif du processus d'innovation	41
1.3.1.2. Rôle des brevets dans le transfert des connaissances	44
1.3.2. De l'importance de l'espace pour innover	47

TABLE DES MATIÈRES

1.3.2.1.	Le rôle secondaire des LKS dans les premiers modèles d'agglomération de P. Krugman (1991a,b)	48
1.3.2.2.	La prise en compte de l'importance des LKS comme facteur d'agglomération	50
1.3.3.	Une discussion de la nature et de l'importance des LKS dans la production d'innovation	51
1.3.3.1.	L'espace n'est pas structuré seulement par les LKS	52
1.3.3.2.	L'innovation est liée aux LKS mais aussi aux DKS et à des proximités non géographiques	52
1.4.	Frontières et canaux de diffusion des connaissances	55
1.4.1.	Les caractéristiques des territoires comme déterminant des frontières de la diffusion des connaissances	55
1.4.1.1.	Les Systèmes Territoriaux d'Innovation (STI)	56
1.4.1.2.	Les effets de proximité dans les systèmes territoriaux d'innovation	60
1.4.2.	Les canaux de transmission des connaissances	61
1.4.2.1.	La mobilité des travailleurs	62
1.4.2.2.	Le spin-off	62
1.4.2.3.	Les réseaux	64
1.4.2.4.	Place des TIC dans le transfert des connaissances	65
1.4.3.	De l'intérêt des pôles de compétitivité dans la transmission des connaissances : une analyse par les réseaux	66
1.4.3.1.	Place de la coopération au sein des pôles de compétitivité	67
1.4.3.2.	Impact de la coopération sur la diffusion des connaissances et sur l'innovation	68
1.4.3.3.	Capacité d'absorption et succès de collaboration	69
1.5.	Conclusion	69
2.	Les outils de mesure de la diffusion spatiale des connaissances : Quelle est la bonne approche empirique ?	71
2.1.	Introduction	71
2.2.	L'approche par les citations de brevets	72
2.2.1.	Principe sous-jacent	73
2.2.2.	Co-localisation entre brevets originels et citations	74

2.2.2.1.	L'approche de Jaffe, Trajtenberg et Henderson (approche JTH) (1993)	76
2.2.2.2.	Approche d'Almeida et Kogut (1997a,b; 1999)	78
2.2.3.	Proximité géographique et flux de connaissance	79
2.2.4.	Limites de l'approche par les citations de brevets	80
2.2.4.1.	Niveau trop agrégé des classes de brevets	80
2.2.4.2.	Les citations ne traduisent pas toujours des <i>spillovers</i>	81
2.2.4.3.	Difficultés de détermination de la localisation	81
2.3.	La Fonction de Production de Connaissances	83
2.3.1.	Mise en évidence des relations Input/output et indice de coïncidence géographique	84
2.3.2.	Prise en compte de la dimension spatiale des externalités de connaissance : le modèle fondateur d'Anselin, Varga et Acs (1997)	86
2.3.2.1.	Une mesure de co-localisation	88
2.3.2.2.	Une mesure de gravité	88
2.3.2.3.	Une mesure de couverture	88
2.3.3.	La dimension technologique des externalités de connaissance	90
2.3.4.	Quelques prolongements	90
2.3.4.1.	Nombre de frontières communes entre unités géographiques	91
2.3.4.2.	Découpage par zone géographique	92
2.4.	Cadre d'analyse empirique et données	95
2.4.1.	Hypothèses et description de l'étude	95
2.4.2.	Données	99
2.4.2.1.	La base de données R&D	100
2.4.2.2.	La base de données brevets	102
2.4.2.3.	Les données sur les pôles de compétitivité fournies par la Direction Générale de la Compétitivité, de l'Industrie et des Services (DGCIS)	104
2.5.	Application de l'approche par cercles concentriques sur données en coupe	105
2.5.1.	Variables et modèle	105
2.5.1.1.	La variable dépendante	105
2.5.1.2.	Les variables indépendantes	107
2.5.1.3.	Le modèle	113
2.5.2.	Estimations et résultats	114

TABLE DES MATIÈRES

2.6.	Application de l'approche concentrique sur données de Panel	120
2.6.1.	Modèle et variables	120
2.6.2.	Estimations et résultats	127
2.7.	Conclusion	133
3.	Diffusion spatiale des connaissances en France : Une analyse exploratoire de données spatiales	135
3.1.	Introduction	135
3.2.	Concentration et agglomération de l'innovation en France	136
3.2.1.	Mesure de la concentration des brevets et des dépenses en R&D en France	136
3.2.2.	En quoi la concentration diffère-t-elle de l'agglomération ?	140
3.3.	Intérêt d'une prise en compte de la dépendance spatiale	142
3.3.1.	Définition et sources de l'autocorrélation spatiale	143
3.3.2.	Les matrices de poids	144
3.3.2.1.	Matrices de contiguïté	145
3.3.2.2.	Les matrices des <i>k-plus proches voisins</i>	146
3.3.2.3.	Les matrices de distance	146
3.3.2.4.	Standardisation des matrices de poids	147
3.3.3.	Comparaison et critère de choix des matrices	148
3.4.	Analyse exploratoire de données spatiales et identification des <i>clusters</i> innovants en France	150
3.4.1.	Analyse globale de l'autocorrélation	152
3.4.1.1.	Statistique <i>I</i> de Moran globale	152
3.4.2.	Analyse locale de l'autocorrélation	156
3.4.2.1.	Diagramme de Moran	157
3.4.2.2.	Statistiques locales et cartes de significativité de Moran	158
3.4.3.	Une application des outils de mesure de l'autocorrélation spatiale locale sur données d'innovation	161
3.4.3.1.	Détection de l'autocorrélation spatiale locale au niveau des brevets	161
3.4.3.2.	Autocorrélation spatiale au niveau des dépenses en R&D	166
3.4.3.3.	Autocorrélation spatiale au niveau des dépenses en R&D externe	168
3.5.	Conclusion	170

4. De l'intérêt d'une analyse économétrique spatiale appliquée à des données de panel	173
4.1. Introduction	173
4.2. Modèles économétriques spatiaux sur données de panel	175
4.2.1. Modèle Spatial Autorégressif (SAR)	176
4.2.2. Modèle avec autocorrélation des erreurs (SEM)	178
4.2.3. Modèle régressif croisé (SDM)	179
4.3. Méthodes d'estimation et règles de choix des modèles spatiaux	179
4.3.1. Méthodes d'estimation des modèles spatiaux sur données de panel . .	180
4.3.1.1. Estimation par le maximum de vraisemblance	181
4.3.2. Règles de choix des modèles spatiaux	183
4.3.2.1. Les tests du multiplicateur de Lagrange	183
4.3.2.2. Test d'Hausman dans le cas spatial	184
4.4. Résultats des estimations des modèles SAR, SEM et SDM	185
4.4.1. Tableaux des résultats	186
4.4.1.1. Résultat d'estimation du modèle SAR	186
4.4.1.2. Résultat d'estimation du modèle SEM	187
4.4.1.3. Résultat d'estimation du modèle SDM	188
4.4.2. Critères du choix du modèle spatial sur données de panel	189
4.4.3. Analyse des résultats et critères du choix du modèle avec effets fixes adéquat	190
4.5. Décomposition de l'effet spatial en effets directs et effets indirects	192
4.5.1. Du modèle SDM aux effets directs et indirects	193
4.5.2. Effets <i>feedbacks</i> ou effets de rétroaction simultanée	196
4.5.3. Résultats de la décomposition en effets directs et indirects	197
4.5.3.1. Effets directs et indirects des variables explicatives	197
4.5.3.2. Répartition des effets directs et indirects par matrice de contiguïté d'ordre 0 à 4	198
4.6. Conclusion	203
Conclusion générale	205
	265

TABLE DES MATIÈRES

Bibliographie	211
Annexes	243
Table des matières	259

Proximité géographique, diffusion des connaissances et innovation : une étude sur les départements français métropolitains

Cette thèse vise à approfondir les recherches théoriques et empiriques sur la diffusion spatiale des connaissances en France métropolitaine. La thèse offre tout d'abord une revue de la littérature économique sur la relation connaissance-externalités-innovation. Elle propose ensuite une analyse empirique utilisant deux approches. La première est l'approche par les cercles concentriques. Nous l'utilisons de deux façons : sur des données en coupe transversale puis sur des données de panel. Les résultats trouvés sont divergents ce qui nous pousse à utiliser une deuxième approche plus précise basée sur les techniques de l'économétrie spatiale. Après avoir décrit de manière détaillée les dynamiques spatiales de l'innovation entre 2002 et 2008, nous estimons des modèles spatiaux (SAR, SEM et SDM) sur nos données de panel. Ensuite, cette deuxième approche propose une décomposition spatiale de l'impact de la R&D en effet direct et effet indirect. Cela permet une mesure plus fine des externalités de connaissance. Les résultats principaux de cette thèse sont que la proximité géographique a des effets positifs sur la productivité de la R&D ; que le contenu de la R&D interne en connaissances tacites lui donne un avantage sur la R&D externe ; que la diversité industrielle des territoires a des effets positifs sur l'innovation ; et que les pôles de compétitivité améliorent la productivité de l'innovation grâce à la richesse des canaux de transmission des connaissances dont ils disposent.

Mots-clés : Diffusion des connaissances, externalités de connaissance, proximité géographique, innovation, économétrie spatiale, clusters, pôle de compétitivité.

Geographical proximity, knowledge diffusion and innovation : a study on French metropolitan regions

The aim of this thesis is to further theoretical and empirical research on the spatial diffusion of knowledge in France. First, this thesis provides a review of the economic literature on the relation knowledge-innovation-externalities. It then offers an empirical analysis using two main approaches. The first one is the concentric circles method. The use of this approach is conducted in two stages : with cross-sectional data and with panel data. The results are divergent which pushed us to use a second more precise approach based on the techniques of spatial econometrics. After comparing the spatial dynamics of innovation between 2002 and 2008, we made estimates using spatial models (SAR, SEM and SDM). Then, we propose in this second approach a spatial decomposition of the R&D impact into a direct and indirect effect. It provides a more precise measure of knowledge spillovers. The main results of this thesis are that geographical proximity has a positive effect on the productivity of R&D ; that internal R&D has higher productivity than external R&D because of its tacit knowledge content ; that industrial diversity of territories has positive effects on innovation ; and that clusters enhance the productivity of innovation through the richness of their knowledge transmission channels.

Keywords : Knowledge diffusion, knowledge spillovers, geographical proximity, innovation, spatial econometrics, clusters.