Titre: Title:	Knowledge flows in clusters and innovation networks : the case of canadian biotechnology and nanotechnology
Auteur: Author:	Andrea Schiffauerova
Date:	2008
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Schiffauerova, A. (2008). Knowledge flows in clusters and innovation networks : the case of canadian biotechnology and nanotechnology [Thèse de doctorat, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/8190/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/8190/
Directeurs de recherche: Advisors:	Catherine Beaudry
Programme: Program:	Non spécifié

UNIVERSITÉ DE MONTRÉAL

KNOWLEDGE FLOWS IN CLUSTERS AND INNOVATION NETWORKS: THE CASE OF CANADIAN BIOTECHNOLOGY AND NANOTECHNOLOGY

ANDREA SCHIFFAUEROVA
DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIAE DOCTOR (PH.D.)
(GÉNIE INDUSTRIEL)
DÉCEMBRE 2008

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-48891-1
Our file Notre référence
ISBN: 978-0-494-48891-1

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée :

KNOWLEDGE FLOWS IN CLUSTERS AND INNOVATION NETWORKS: THE CASE OF CANADIAN BIOTECHNOLOGY AND NANOTECHNOLOGY

présentée par : SCHIFFAUEROVA Andrea
en vue de l'obtention du diplôme de : Philosophiae Doctor
a été dûment acceptée par le jury d'examen constitué de :

Mme DE MARCELLIS-WARIN Nathalie, Doctorat, présidente
Mme BEAUDRY Catherine, D. Phil., membre et directrice de recherche
Mme LEFEBVRE Élisabeth, Ph. D., membre
M. NIOSI Jorge, Doct. Soc., membre externe

ACKNOWLEDGEMENTS

I owe a sincere gratitude to Prof. Catherine Beaudry for her invaluable guidance and advice which was of paramount importance in making this thesis possible.

I would also like to express special thanks to Ahmad Barirani for the creation of the extraction program and to Prof. Martin Trépanier for his help with the databases. Also I would like to thank the personnel of the Department of Mathematics and Industrial Engineering, especially to Diane Bernier and Suzanne Guindon.

Sincere thanks to all my friends and my brother for their support. I do not think I could ever pay back my parents for their endless love and encouragement. Finally, many special uncountable thanks and appreciation to Mamoun, my dear husband, for his continuous support, immense encouragement, patience and love.

RÉSUMÉ

Les objectifs de recherche de cette thèse sont organisés autour de quatre thèmes principaux. Le premier thème concerne le débat sur la performance économique des compagnies dans les régions spécialisées ou diversifiées. Ici, le but principal est de fournir une taxonomie des articles scientifiques et d'étudier les raisons derrière cette contradiction. Le deuxième thème de la thèse concerne la création de l'innovation dans les grappes canadiennes de biotechnologie. L'objectif est d'identifier, d'analyser et de caractériser ces grappes, avec une concentration spéciale sur la politique de propriété intellectuelle dans les établissements canadiens. La collaboration et les réseaux d'innovation sont le troisième sujet principal de ce travail. Ici la thèse cherche d'abord à examiner les aspects géographiques de la collaboration et de l'impact de la proximité géographique sur le choix des partenaires de collaboration en biotechnologie et nanotechnologie. Ensuite, le but est de comparer les caractéristiques des réseaux de collaboration de biotechnologie et de nanotechnologie, et de souligner leurs rôles dans la diffusion efficace de la connaissance. Finalement, le quatrième thème traite des inventeurs proéminents dans les grappes canadiennes de biotechnologie et de leurs rôles dans les réseaux d'innovation. Un premier objectif ici est de développer des méthodologies innovatrices pour identifier les individus clés dans le réseau de collaboration. Un second objectif concerne les scientifiques étoiles. La thèse envisage de trouver différents moyens pour leur identification et d'étudier leurs positions géographiques aussi bien que leurs positions dans le réseau.

L'approche principale de la thèse consiste en l'exploitation d'information extraite à partir de la base de données de brevet du Bureau des brevets et de marques déposées des États-Unis (USPTO). L'information est utilisée pour la description de la création de l'innovation en biotechnologie et en nanotechnologie au Canada aussi bien que pour la création des réseaux d'innovation basés sur les liens parmi les co-inventeurs. Les méthodes d'analyse de réseaux sociaux sont employées afin de créer, de caractériser et d'évaluer ces réseaux d'innovation.

Cette thèse apporte une contribution importante en clarifiant pourquoi les résultats des études qui concernent la question de l'urbanisation et de la spécialisation d'une région sont souvent contradictoires. Les contradictions peuvent être expliquées par la puissance variée des forces d'agglomération à travers des industries, pays ou périodes de temps, mais également par les raisons méthodologiques et les divers indicateurs des externalités de MAR et de Jacobs utilisées dans la recherche. Généralement, cette thèse suggère que dans les régions avec des industries matures et de basse technologie, la politique régionale devrait soutenir le développement d'un ensemble étroit d'activités économiques dans la région, ce qui devrait mener à une plus grande productivité. Mais dans les régions de pointe, la politique devrait se concentrer sur la création d'un ensemble divers d'activités économiques, ce qui devrait augmenter le développement économique.

On trouve que l'activité innovatrice au Canada est concentrée dans plusieurs endroits qui correspondent plus ou moins aux zones métropolitaines principales. Ces grappes ont été décrites par la quantité et la qualité de brevets, par la nature des activités de biotechnologie, et par les caractéristiques de propriétaires de brevets et leur propension à collaborer. Environ la moitié des brevets sont possédés par des entreprises. Cependant, la recherche financée par des ressources publiques est très importante pour le secteur de biotechnologie au Canada. Les universités sont les établissements les plus actifs en biotechnologie et elles sont aussi les plus grands producteurs des brevets. La contribution des laboratoires gouvernementaux à la recherche et développement en biotechnologie est également substantielle.

La majorité des collaborations des inventeurs canadiens se réalise au sein de grappes canadiennes, alors que la collaboration parmi les inventeurs qui résident dans les grappes canadiennes différentes est beaucoup moins commune. Les liens internationaux forment la proportion la plus élevée parmi toutes les collaborations en dehors des grappes et les partenaires de collaboration étrangers les plus populaires résident aux États-Unis. La présence d'inventeurs étrangers est critique pour la transmission de la connaissance entre les inventeurs canadiens eux-mêmes. Les étrangers sont extrêmement importants, parce

qu'ils relient les inventeurs canadiens de grappes différentes (ou même ceux des mêmes grappes) les uns aux autres.

La thèse suggère que la distance géographique joue un rôle très important. Les inventeurs canadiens préfèrent collaborer avec des partenaires locaux ou relativement proches. Néanmoins, si les collaborateurs nécessaires et adéquats ne sont pas trouvés à l'intérieur d'une distance d'environ 600 kilomètres, l'importance du facteur géographique diminue considérablement, puisque dans ce cas-ci les inventeurs optent souvent pour des partenaires de coopération très éloignés ou même outre-mer.

Les structures de collaboration dans les réseaux de biotechnologie et de nanotechnologie sont tout à fait distinctes. Le réseau d'innovation de biotechnologie est plus grand, plus développé et moins fragmenté que celui de nanotechnologie. La fragmentation plus élevée du réseau de nanotechnologie est expliquée par une plus grande disparité parmi les domaines en nanotechnologie comparée aux spécialisations de biotechnologie qui sont beaucoup plus étroitement liés.

L'architecture du réseau des inventeurs canadiens de biotechnologie a été étudiée dans le cadre de deux concepts différents: D'abord, la collaboration parmi des inventeurs travaillant dans les mêmes grappes (espace géographique); ensuite, la coopération parmi des inventeurs qui sont directement ou indirectement interconnectés dans des composantes de réseau (espace technologique). Les deux espaces de collaboration se chevauchent dans une certaine mesure, mais ils diffèrent dans leurs structures. Les portiers de la connaissance sont les inventeurs qui font le pont entre ces deux espaces et permettent ainsi d'alimenter les grappes de biotechnologie avec la nouvelle connaissance extérieure à la grappe. Cette thèse propose des indicateurs, qui mesurent l'importance de chaque inventeur en tant que fournisseur de connaissance externe pour la grappe (ou pour le Canada). Seulement environ 10%-20% de tous les inventeurs canadiens dans la plupart des grappes ont été identifiés comme portiers qui sont responsables de l'apport d'information externe à la grappe.

Afin d'identifier les scientifiques étoiles dans les grappes canadiennes de biotechnologie, nous avons proposé de nouvelles mesures basées sur la quantité et la qualité de brevets et sur le nombre de citations des articles scientifiques. La majorité des scientifiques étoiles ont également été identifiés comme portiers responsables de l'apport d'information externe qui contribue fortement au potentiel innovateur canadien.

Par la caractérisation des grappes canadiennes de pointe et par l'éclaircissement du processus de transmission de la connaissance par des réseaux d'innovation, cette thèse contribue non seulement à la compréhension des transferts de connaissance dans les secteurs de pointe, mais aussi à la facilitation de l'innovation dans les grappes canadiennes. Les résultats de cette recherche ont été présentés à diverses conférences et sont considérés pour la publication dans plusieurs journaux et livres.

ABSTRACT

The research objectives of this thesis are organized around four main themes. The first theme concerns the debate on the economic performance and growth of firms in the specialized versus diversified clusters or regions, which has yet to reach conclusive results. Here, the main aim is to provide taxonomy of scientific articles and to investigate the reasons behind this inconsistency. The second theme of the thesis concerns the creation of innovation in Canadian biotechnology clusters. The objective is to identify, analyze and characterize these clusters, with a special focus on the intellectual property politics in Canadian institutions. The innovation networks and collaboration are the third major topic of this work. Here the thesis first seeks to examine the geographical aspects of the collaboration and the impact of geographical proximity on the selection of the collaboration partners in Canadian biotechnology and nanotechnology. Afterwards, the goal is to compare the characteristics of the biotechnology and nanotechnology collaboration networks and to highlight their role in the efficient knowledge diffusion. Finally, the fourth theme deals with the prominent inventors in Canadian biotechnology clusters and their roles in the innovation networks. One objective here is to develop innovative methodologies to identify the key individuals (knowledge gatekeepers) in the collaboration network. Another objective concerns the star scientists. The thesis intends to find different ways for their identification and to investigate their geographical and network positions.

The main approach of the thesis consists in the creative exploitation of the large amounts of information extracted from the patent database of the United States Patent and Trademark Office (USPTO). The information is used to describe the creation of biotechnology and nanotechnology innovation in Canada and to build the innovation networks based on the patent co-inventorship links. The methods of social network analysis are used to create, characterize and evaluate these innovation networks.

The thesis has made a major contribution in clarifying why the results of the studies concerning the urbanization and localization issue are often conflicting. The inconsistency

is explained not only the by differences in the strength of agglomeration forces across industries, countries or time periods, but also by methodological issues and the various indicators of MAR and Jacobs externalities used in the research. In general, this thesis suggests that in regions with mature, low tech industries, regional policy should emphasize the development of a narrow set of economic activities in the region, which will presumably lead to greater productivity. In high tech regions, on the other hand, policy should focus on the creation of a diverse set of economic activities, which should enhance economic development.

It was shown that innovative activity in Canada is concentrated in several locations which roughly correspond to the larger metropolitan areas. The thesis has made a contribution by making a profile description for the Canadian biotechnology clusters in terms of patenting quality and quantity, the nature of biotechnology activities, the properties of assignees and their propensity to collaborate. Around half of the patents are assigned to firms. However, publicly-funded research is highly important for biotechnology in Canada. Universities are the most active institutions in biotechnology and the greatest producers of patents. The contribution of the government laboratories to the biotechnology research and development is also substantial.

Most of the collaborative activity of Canadian inventors takes place within Canadian clusters, while the inter-cluster collaboration in Canada is much less common for both biotechnology and nanotechnology. International ties account for the highest proportion of all the collaborations outside the clusters and the most popular foreign collaboration partners for Canadian inventors reside in the USA. The presence of foreigners in the Canadian collaboration network is critical for the transmission of knowledge between Canadian inventors themselves. Foreigners are extremely important in connecting Canadian inventors from different clusters (or even those from the same cluster) together.

This thesis suggests that the distance plays an important role in selecting the research collaborators in both biotechnology and nanotechnology. An overwhelming preference of the Canadian inventors is towards local and relatively proximate partnerships. Nonetheless, if the suitable collaborators are not found within the distance of around 600 km, the importance of the geographical factor significantly decreases, since in this case both biotechnology and nanotechnology inventors quite often opt for very distant or overseas cooperation partners.

The collaborative structures within biotechnology and nanotechnology networks are quite distinct. The biotechnology innovation network is larger, more developed and less fragmented than that of nanotechnology. The higher fragmentation of the nanotechnology network is explained by the greater disparity among the nanotechnology specializations compared to the more closely related biotechnology fields.

The architecture of the network of Canadian biotechnology inventors was investigated within two different concepts: First, collaboration among inventors working in clusters (geographical proximity); second, cooperation among inventors who are directly or indirectly interconnected in network components (cognitive proximity). The geographical and technological dimensions both nurture the growth of the cluster and promote innovation through a dynamic interaction of the actors localized in clusters who absorb external knowledge through the local and non-local networks. The geographical and technological collaboration spaces thus overlap to a certain extent, but they differ in their structures. Gatekeepers are the inventors who bridge over the geographical and technological spaces and hence enable the nurturing of biotechnology clusters with fresh external knowledge. This thesis proposes indicators, which measure each inventor's importance as a procurer of external knowledge for the cluster (or for Canada) based on the share of innovative production to which he thereby contributes. Only around 10%-20% of all inventors in most clusters were identified as gatekeepers who are responsible for the inflow of external information to the cluster.

Star scientists are recognized as a key driving force behind the growth and innovation in biotechnology. In order to identify the most prolific inventors in Canadian biotechnology clusters, new measures based on the patent quantity, quality and the number of forward citations in scientific articles were proposed. The majority of the star scientists so defined are also identified as gatekeepers responsible for the inflow of external information which highly contributes to the Canadian innovative potential.

By characterizing Canadian high tech clusters and shedding light on the knowledge transmission processes that are carried out through innovation networks, this thesis has contributed to the understanding of knowledge transfers that characterise high technology sectors in Canada. The results of this research have been presented in well-recognized conferences and are also being under consideration for publication in several peer-reviewed journals and books.

CONDENSÉ EN FRANÇAIS

Les flux de connaissance ont été identifiés comme étant le facteur explicatif principal pour la création des grappes géographiques des entreprises innovatrices et aussi comme un élément critique de la contribution des grappes à la croissance économique régionale. Les distances géographiques et cognitives entre les inventeurs, les structures de réseaux de collaboration et les positions des inventeurs dans le réseau ont un effet important sur la diffusion de la connaissance par l'intermédiaire du réseau, et par conséquent sur les résultats économiques et la propension d'innover des firmes au sein des grappes. Donc, afin de comprendre les éléments clés qui soutiennent la croissance des grappes industrielles au Canada, il est nécessaire de bien comprendre la diffusion de la connaissance dans les grappes et dans les réseaux d'innovation. Cette thèse examinera ces thèmes en détail.

Questions de recherche

Les grappes se trouvent au cœur de la thèse. La première question étudiée concerne le rôle de la composition industrielle d'une grappe: Les grappes peuvent être spécialisées, ce qui est le cas quand la plupart des entreprises et des services de soutien de la grappe appartenant à une industrie principale; mais les grappes peuvent être également diversifiées, ce qui est le cas quand beaucoup d'industries diverses sont représentées dans une grappe. La question de la composition des activités économiques dans la grappe et de son influence sur la croissance économique de la région a été déjà posée par beaucoup de chercheurs. Étant donné que leurs conclusions ont été très contradictoires, cette question est abordée dans cette thèse à nouveau. Cependant, ce travail se démarque des questions habituelles (Est-ce que les entreprises ont de meilleurs résultats économiques dans les grappes spécialisées ou diversifiées? Est-ce la spécialisation ou la diversité qui favorise le plus l'innovation?, etc.) et explore la raison pour laquelle la communauté académique n'a pas encore établi un consensus.

Néanmoins, la thèse se concentre spécifiquement sur les grappes canadiennes de biotechnologie. Les questions principales de recherche posées ici sont: Est-ce que l'innovation canadienne en biotechnologie se réalise principalement dans les grappes? Où sont exactement ces grappes? Qui sont les inventeurs dans les grappes canadiennes de biotechnologie? Quel est le rôle des compagnies, des universités et des laboratoires gouvernementaux dans le processus innovateur de biotechnologie canadienne? Est-ce que ces compagnies ou établissements collaborent en créant des innovations? Qui sont leurs partenaires principaux de coopération et où résident-ils? Est-ce que la collaboration se réalise principalement à l'intérieur de la grappe, entre les grappes ou est-elle plutôt internationale? Est-ce que la distance géographique joue un rôle dans le choix du partenaire de collaboration?

Les autres questions concernent les chercheurs de biotechnologie pris individuellement: Est-ce que tous les chercheurs de la grappe ont la même importance pour la communication de la connaissance obtenue hors de la grappe? Est-ce que il y a des chercheurs qui sont instrumentaux à l'alimentation des grappes avec de nouvelles informations venant de l'extérieur? Comment on peut identifier ces portiers de la connaissance et comment peut-on évaluer leur importance? Est-ce que il y a quelques inventeurs qui produisent considérablement plus d'innovation que d'autres? Comment identifier ces scientifiques étoiles? Quelles sont leurs positions dans le réseau de collaboration? Est-ce que ces scientifiques étoiles sont aussi les portiers de la connaissance ou les deux rôles sont séparés?

Toutes ces questions se rapportent à grappes canadiennes de biotechnologie. Cependant, la thèse offre également une comparaison avec une autre technologie de pointe qui est aussi très importante pour le Canada – la nanotechnologie. Ici, on pose questions suivantes: Est-ce que l'évolution de l'innovation en biotechnologie et en nanotechnologie au Canada est semblable? Est-ce que les modèles de collaboration sont similaires dans les deux domaines? Est-que les inventeurs ont les préférences semblables ou distinctes dans le choix de leurs partenaires de collaboration? Est-ce que les réseaux de la connaissance sont comparables? Quelles sont les différences et les similitudes entre ces deux domaines?

La section suivante formule les objectifs principaux qui permettent répondre aux questions posées ci-dessus.

Objectifs de recherche et leur accomplissement

La thèse traite d'abord la question des externalités de la connaissance et explore leur rôle dans les régions ou les grappes. On considère deux types d'externalités, qui jouent un rôle important dans le processus de création et diffusion de la connaissance: les externalités de spécialisation (Marshall-Arrow-Romer ou MAR) qui agissent principalement dans une industrie spécifique, et les externalités de diversité (Jacobs) qui agissent entre les secteurs. Par conséquent, la performance économique de la grappe est soit favorisée par la concentration d'une industrie particulière dans une grappe (MAR) ou c'est la diversité des industries dans une région qui favorise la croissance et l'innovation (Jacobs). La question de savoir si la spécialisation ou la diversité (l'urbanisation) des activités économiques favorise mieux le développement dans la région a été le sujet d'un débat passionné dans la littérature économique. Il y a une grande contradiction dans les résultats des travaux de recherche qui fournissent l'évidence pour l'appui ou pour l'opposition à l'une ou l'autre de ces deux théories. Le premier objectif de la thèse est donc d'étudier les raisons derrière les résultats contradictoires de la littérature en ce qui concerne l'impact des externalités de spécialisation et d'urbanisation sur la performance économique des entreprises dans les régions et les grappes. La thèse offre un recensement des articles qui ont traité le sujet et examine les similitudes entre ces diverses études.

Le deuxième objectif est de décrire la création de l'innovation dans les grappes canadiennes de biotechnologie. La thèse identifie, analyse et caractérise les grappes canadiennes de biotechnologie avec une concentration spéciale sur la quantité et la qualité de brevets, sur la nature des activités de biotechnologie et sur les propriétaires des brevets et leur propension à collaborer. Le rôle critique de la recherche publique canadienne en biotechnologie est examiné et l'importance de la propriété intellectuelle et des bureaux de transfert de technologie aux universités canadiennes est identifiée.

La thèse aborde ensuite la question des grappes canadiennes de biotechnologie et se concentre sur la collaboration au sein des grappes et entre celles-ci. Plus spécifiquement, le troisième objectif consiste à étudier le rôle de la géographie dans la collaboration. Le modèle de collaboration dans l'innovation canadienne en biotechnologie est décrit, puis l'importance des circonstances et de la proximité géographiques dans le choix des partenaires de collaboration est examinée.

Le quatrième objectif consiste à comparer l'innovation en biotechnologie et en nanotechnologie au Canada. La comparaison de l'évolution des brevets de biotechnologie et de nanotechnologie et des caractéristiques principales de collaboration dans les grappes de biotechnologie et de nanotechnologie est faite. Le but principal ici est cependant la recherche sur la collaboration locale dans les sous-réseaux (basés sur les grappes). Les propriétés structurales des sous-réseaux de biotechnologie et de nanotechnologie sont examinées et comparées, et l'efficacité des sous-réseaux dans la diffusion de la connaissance et dans la création d'innovation est discuté.

Les deux derniers objectifs demeurent dans le domaine de l'innovation canadienne en biotechnologie, mais on se concentre maintenant sur les inventeurs, plus spécifiquement sur les individus principaux dans le processus d'innovation. Le cinquième objectif est d'étudier des portiers de la connaissance dans le réseau de collaboration. Les portiers sont les inventeurs qui font le pont entre deux espaces de collaboration - l'espace géographique (grappes) et l'espace technologique (réseau) - et qui permettent ainsi l'alimentation des grappes avec la connaissance externe. Nous proposons une méthode qui facilite l'identification des portiers dans le réseau de collaboration. Le rôle des portiers pour les grappes et pour Canada est alors discuté.

Le sixième objectif consiste à identifier les scientifiques étoiles canadiens en biotechnologie et à examiner leur rôle dans le réseau d'innovation. D'abord, nous proposons une nouvelle méthode pour l'identification de scientifiques étoiles. Ces inventeurs sont identifiés en utilisant la quantité de leurs brevets (scientifiques étoiles), la quantité et la qualité de ces brevets (QQ-scientifiques étoiles), et la quantité de citations des articles scientifiques. D'ailleurs, les positions de ces scientifiques dans le réseau de collaboration sont étudiées. La thèse examine également le chevauchement entre les portiers de la connaissance et les scientifiques étoiles ou les QQ-scientifiques étoiles.

Données et méthodologie

L'approche principale consiste en l'exploitation de l'information contenue dans les bases de données de brevets en biotechnologie et en nanotechnologie. La base de données de brevets en biotechnologie utilisée pour l'analyse empirique vient de la base de données du Bureau des brevets et de marques déposées des États-Unis (USPTO). Une des tâches initiales était donc de choisir une définition précise et pratique de la biotechnologie, qui permettrait l'identification des brevets appropriés en biotechnologie dans l'USPTO. Un programme automatisé d'extraction a été alors employé pour collecter les informations exigées des brevets en biotechnologie. La base de données finale contient tous les brevets dans lesquels au moins un inventeur réside au Canada, et elle comporte 3550 brevets.

Afin d'établir le réseau des inventeurs canadiens de nanotechnologie, les données des brevets contenues dans la banque de données de Nanobank ont été employées. Nanobank est une bibliothèque numérique publique comportant des données sur des articles scientifiques, des brevets et des subventions fédérales en nanotechnologie. La base de données de brevets de Nanobank s'appuie aussi sur les données des brevets de l'USPTO. Nous avons également choisi seulement les brevets dans lesquels au moins un inventeur réside au Canada. D'ailleurs, des filtres additionnels ont été utilisés afin de sélectionner les brevets qui sont strictement liés à la nanotechnologie. La base de données canadienne de brevets de nanotechnologie ainsi créée comporte 1443 brevets.

L'information contenue dans ces deux bases de données a été utilisée pour l'analyse et la caractérisation des grappes canadiennes. Après, les réseaux d'innovation de biotechnologie et de nanotechnologie ont été créés à partir de ces deux bases de données, en traçant tous les liens parmi les co-inventeurs de chaque brevet particulier. Le concept de l'analyse de réseau social a été utilisé pour créer des connexions entre les inventeurs et le programme d'analyse de réseau social PAJEK a été employé pour la construction de réseaux d'innovation. Une analyse détaillée de ces réseaux a permis la description de leurs propriétés structurales et a facilité la compréhension du comportement de collaboration des inventeurs à l'intérieur ou à l'extérieur des grappes canadiennes de pointe. La section suivante décrit les résultats obtenus et les conclusions.

Résultats et conclusions

Cette thèse a analysé une grande gamme des études démontrant l'impact positif des externalités de Marshall et de Jacobs sur la performance régionale. De plus, un nombre non négligeable d'effets négatifs de MAR a été observé, ce qui implique que la spécialisation d'une région peut en fait aussi en gêner la croissance économique. Il est beaucoup moins probable que la diversification produise cet impact négatif. La thèse a apporté une contribution importante en clarifiant les raisons pour lesquelles les résultats de ces études sont souvent contradictoires et en spécifiant ce qui importe, et comment ceci fait la différence. Les contradictions peuvent être expliquées par la puissance variée des forces d'agglomération à travers des industries, pays ou périodes de temps, mais également par les raisons méthodologiques et les divers indicateurs des externalités de MAR et de Jacobs utilisées dans la recherche.

Les avantages des régions spécialisées ou diversifiées pour des industries particulières ont été évalués dans une analyse plus détaillée des secteurs industriels. On trouve que dans les secteurs de basse technologie, les externalités de Marshall ont des effets plus forts que les externalités de Jacobs. La situation dans les secteurs de moyenne technologie donne des résultats similaires pour les deux théories, mais diffère pour les secteurs de pointe pour lesquels le développement est légèrement favorisé dans les régions diversifiés, alors que les effets des externalités de Marshall sont moins prononcés. La diversification est également un instigateur de croissance dans les services. De plus, le rôle des externalités varie selon la maturité de l'industrie. Les externalités de Jacobs prédominent pendant les étapes initiales du cycle de vie de l'industrie, tandis que les externalités de Marshall entrent plus tard, et à la fin, la spécialisation va en fait gêner la croissance économique.

Les implications de cette recherche pour la politique publique sont tout à fait importantes. Généralement, cette thèse suggère que dans les régions avec des industries matures et de basse technologie, la politique régionale devrait soutenir le développement d'un ensemble étroit d'activités économiques dans la région, ce qui devrait mener à une plus grande productivité. Mais dans les régions de haute technologie, la politique devrait se concentrer sur la création d'un ensemble divers d'activités économiques, ce qui devrait augmenter le développement économique. Cependant, étant donné que les avis académiques sont tellement contrastants et leurs conclusions souvent contradictoires, la politique de développement régional qui soutient ou discrimine certaines activités industrielles ou certaines technologies devrait être appliquée avec prudence, tout au moins jusqu'à ce que cette problématique soit entièrement clarifiée.

Le sujet principal de la thèse porte sur les grappes canadiennes de pointe. On trouve que l'activité innovatrice au Canada est concentrée dans plusieurs endroits qui correspondent plus ou moins aux zones métropolitaines principales: 12 grappes de biotechnologie et 8 grappes de nanotechnologie ont été identifiées. En biotechnologie, plus de la moitié de tous les inventeurs canadiens résident dans les trois plus grandes grappes - Toronto, Montréal et Vancouver. En nanotechnologie, c'est principalement la grappe de Toronto qui domine le secteur industriel, parce qu'environ un quart de tous les inventeurs canadiens de nanotechnologie y vivent. Ces grappes ont été décrites en se concentrant sur la quantité et la qualité de brevets, sur la nature des activités de biotechnologie, et sur les caractéristiques de propriétaires de brevets et leur propension de collaborer.

Environ la moitié des brevets sont possédés par des entreprises. Cependant, la recherche financée par des ressources publiques est très importante pour le secteur de la biotechnologie au Canada. Les universités sont les établissements les plus actifs en biotechnologie et elles sont aussi les plus grands producteurs de brevets. La production des brevets pourtant diffère énormément parmi les universités canadiennes: plusieurs universités renommées qui sont très actives dans la recherche en biotechnologie possèdent un nombre de brevets très inférieur à d'autres universités moins actives. On a expliqué ces différences par l'existence, la qualité et l'efficacité du bureau du transfert technologique disponible à ces universités, et par les règles et les politiques des universités concernant la propriété intellectuelle. La contribution des laboratoires gouvernementaux à la recherche et développement en biotechnologie est également substantielle. Dans les grappes qui accueillent les cinq instituts du Conseil national de recherches Canada (Montréal, Vancouver, Ottawa, Saskatoon et Halifax), les sous-réseaux (basés sur les grappes) de biotechnologie sont généralement mieux développés et la recherche y est mieux organisée. Cependant, en nanotechnologie, seulement deux pôles sont présents, Toronto et Edmonton, ce dernier étant toujours en émergence.

Il y a de grandes capacités d'innovation parmi les chercheurs canadiens, mais beaucoup de la propriété intellectuelle en fait quitte le pays. C'est particulièrement évident en nanotechnologie. Presque la moitié de toutes les innovations inventées ou co-inventées par les inventeurs canadiens de nanotechnologie sont possédées par des étrangers. Bien que ce soit les Canadiens qui fassent la recherche, le fruit de leur travail n'est pas approprié par des intérêts canadiens.

Peu d'évidence concernant la coopération des compagnies ou des établissements de biotechnologie et de nanotechnologie a été trouvée. Il y a très peu de coopération à l'intérieur des grappes et encore moins de coopération entre les grappes. Le partenaire de collaboration le plus fréquent pour un établissement ou une compagnie canadiens est un autre établissement ou compagnie à l'étranger (principalement aux États-Unis).

Beaucoup plus de collaboration a été détectée parmi les inventeurs individuels de biotechnologie et de nanotechnologie. La majorité de la collaboration des inventeurs canadiens se réalise dans les grappes canadiennes, alors que la collaboration entre les inventeurs qui résident dans les grappes canadiennes différentes est beaucoup moins commune. Les liens internationaux forment la proportion la plus élevée parmi toutes les collaborations en dehors des grappes, tandis que les partenaires de collaboration étrangers les plus populaires résident aussi aux États-Unis. Environ un tiers des inventeurs identifiés dans les deux bases de données résident à l'étranger. Ces inventeurs sont tellement mêlés au réseau de collaboration des Canadiens que leur présence est en fait absolument critique pour la transmission de la connaissance entre les inventeurs canadiens eux-mêmes. Les étrangers sont extrêmement importants, parce qu'ils relient les inventeurs canadiens de grappes différentes (ou même ceux des mêmes grappes) les uns aux autres.

Quand les inventeurs en biotechnologie et en nanotechnologie sélectionnent leurs collaborateurs de recherches, la distance géographique joue un rôle très important. Les inventeurs canadiens préfèrent collaborer avec les partenaires locaux ou relativement proches. Néanmoins, si les collaborateurs nécessaires et adéquats ne sont pas trouvés à l'intérieur d'une distance d'environ 600 kilomètres, l'importance du facteur géographique diminue considérablement, puisque dans ce cas-ci les inventeurs optent souvent pour des partenaires de coopération très éloignés ou même outre-mer.

Les structures de collaboration dans des réseaux de biotechnologie et de nanotechnologie sont tout à fait distinctes. Le réseau d'innovation de biotechnologie est plus grand, plus développé et moins fragmenté que celui de nanotechnologie. La fragmentation plus élevée du réseau de nanotechnologie est expliquée par une plus grande disparité parmi les domaines en nanotechnologie comparée aux spécialisations de biotechnologie qui sont beaucoup plus étroitement liées.

L'architecture du réseau des inventeurs canadiens de biotechnologie a été étudiée dans le cadre de deux concepts différents: D'abord, la collaboration parmi des inventeurs travaillant dans les mêmes grappes (proximité géographique); ensuite, la coopération parmi des inventeurs qui sont directement ou indirectement interconnectés dans des composants de réseau (proximité cognitive). L'espace géographique (basé sur des grappes) et l'espace technologique (basé sur le réseau) sont tous deux très importants pour la création et la diffusion de la connaissance. Les dimensions géographiques et technologiques soutiennent la croissance de la grappe et favorisent l'innovation par une interaction dynamique des acteurs qui sont localisés dans les grappes et qui absorbent la connaissance externe par les réseaux locaux et non-locaux. Les espaces géographiques et technologiques de collaboration se chevauchent dans une certaine mesure, mais ils diffèrent dans leurs structures. Les points d'interaction entre ces deux espaces de collaboration sont les inventeurs qui sont très bien interconnectés à l'intérieur et à l'extérieur des grappes. On les appelle portiers de la connaissance.

Les portiers sont les inventeurs qui font le pont entre l'espace géographique et l'espace technologique et permettent ainsi d'alimenter les grappes de biotechnologie avec de nouvelles connaissances externes à la grappe ou au Canada. Cette thèse propose des indicateurs, qui mesurent l'importance de chaque inventeur en tant que fournisseur de connaissance externe pour la grappe (ou pour le Canada) et qui sont basés sur la portion de la production des inventions à laquelle il contribue ainsi. Seulement environ 10%-20% de tous les inventeurs canadiens dans la plupart des grappes ont été identifiés comme portiers qui sont responsables de l'apport d'information externe à la grappe.

Les scientifiques étoiles sont reconnus comme une force principale derrière la croissance et l'innovation en biotechnologie. Afin d'identifier les inventeurs les plus proéminents dans les grappes canadiennes de biotechnologie, on a proposé de nouvelles mesures. Celles-ci considèrent seulement la quantité de brevets (inventeurs étoiles), la quantité et la qualité de brevet (QQ- inventeurs étoiles), ou le nombre de citations des articles scientifiques (scientifiques fortement cités). Ces critères ont alors permis de distinguer et de comparer les divers inventeurs proéminents, avec quelques conclusions intéressantes: Les inventeurs étoiles de biotechnologie n'inventent ou ne co-inventent pas nécessairement des brevets de valeur élevée. De plus, les chercheurs et scientifiques fortement cités, qui sont considérés supérieurs dans le domaine de biotechnologie au Canada, ne produisent pas toujours des brevets ou ne les enregistrent pas à l'USPTO.

Finalement, la coïncidence des inventeurs proéminents avec les portiers de la connaissance a été examinée. La grande majorité des inventeurs étoiles, des QQ-inventeurs étoiles et presque la moitié de tous les scientifiques fortement cités ont été également identifiés comme portiers responsables de l'apport d'information externe qui contribue fortement au potentiel innovateur canadien.

Contribution

Le sujet de cette thèse est de grande importance pour le Canada, parce que la thèse se concentre sur deux de ses domaines les plus dynamiques – la biotechnologie et la nanotechnologie. Ces domaines représentent une contribution considérable à l'avancement de la science et l'innovation, ils fournissent des milliers de travaux, aussi bien que de grandes exportations. Par la caractérisation des grappes canadiennes de pointe et par l'éclaircissement du processus de transmission de la connaissance par des réseaux d'innovation, cette thèse contribue non seulement à la compréhension des transferts de connaissance dans les secteurs de pointe, mais aussi à la facilitation de l'innovation dans les grappes canadiennes.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iv
RÉSUMÉ v
ABSTRACT ix
CONDENSÉ EN FRANÇAIS xiii
TABLE OF CONTENTS xxv
LIST OF FIGURES xxix
LIST OF TABLES xxxi
LIST OF APPENDICES xxxv

INTRODUCTION 1

CHAPTER 1 LITERATURE REVIEW 6

1.1 Clusters 7
  1.1.1 Clustering benefits and costs 8
  1.1.2 Performance of the firm in cluster 10

1.2 Localized knowledge spillovers 13
  1.2.1 The concept of the knowledge production function 13
  1.2.2 Localization effects of knowledge spillovers 15
  1.2.3 Knowledge flows from public science 18
  1.2.4 Knowledge spillover mechanisms 28
  1.2.5 Critique of the localized knowledge spillovers theory 39
  1.2.6 Diversification and specialization externalities 44

1.3 Innovation networks 46
  1.3.1 Collective invention 46
  1.3.2 Network structure analysis 47
  1.3.3 Brokers and gatekeepers 55

CHAPTER 2 RESEARCH QUESTIONS, OBJECTIVES AND METHODOLOGY 59

2.1 Research questions 59

2.2 Research objectives 61

2.3 General organization of the thesis 62

2.4 Methodology 64
  2.4.1 Patents and their use in innovation research 64
  2.4.2 Biotechnology database 67
  2.4.3 Nanotechnology database 70
  2.4.4 Social network analysis 76

CHAPTER 3 THE LOCALISATION VERSUS URBANISATION CONTROVERSY 77

3.1 Basic results 78

3.2 Indicators for Marshall and Jacobs externalities 81
  3.2.1 Marshall externalities indicators 82
  3.2.2 Jacobs externalities indicators 86

3.3 Performance measures 89
  3.3.1 Economic growth 91
  3.3.2 Productivity 93
  3.3.3 Innovation 94

3.4 Other circumstances and conditions of analysis 96
  3.4.1 Regional and firm level analyses 97
  3.4.2 Industry range and industrial sectors 99
  3.4.3 Geographical unit 106
  3.4.4 Countries and regions 108
  3.4.5 Period of observation 111
  3.4.6 Size of the firms 112

3.5 Conclusions 112

CHAPTER 4 INNOVATION IN CANADIAN BIOTECHNOLOGY CLUSTERS 116

4.1 Canadian biotechnology clusters 118

4.2 Patent ownership in Canadian biotechnology clusters 120

4.3 Biotechnology field specialization in clusters 131

4.4 Collaboration in Canadian biotechnology based on patent co-assignment 135

4.5 Conclusions 137

CHAPTER 5 GEOGRAPHICAL ASPECTS OF COLLABORATION 141

5.1 International collaborations 144

5.2 Inter-cluster Collaborations 147

5.3 Distance-based analysis of all out-of-cluster collaborations 151

5.4 Conclusions 153

CHAPTER 6 COMPARISON OF BIOTECHNOLOGY AND NANOTECHNOLOGY INNOVATION NETWORKS 156

6.1 Canadian biotechnology and nanotechnology clusters 156

6.2 Collaboration patterns in Canadian biotechnology and nanotechnology 160

6.3 Local collaboration in the cluster-based subnetworks 164

  6.3.1 Collaboration characteristics in the subnetworks 165

  6.3.2 Fragmentation of the subnetworks 167

  6.3.3 Structural cohesion of the subnetworks 171

  6.3.4 Cliquishness in the subnetworks 172

  6.3.5 Centrality of vertices 174

  6.3.6 Centralization of the subnetworks 176

6.3.7 Geodesic distances in the subnetworks 177
6.3.8 Summary of the network properties 179

6.4 Conclusions 181

CHAPTER 7 GATEKEEPERS OF CANADIAN BIOTECHNOLOGY CLUSTERS 185

7.1 Geographical and cognitive proximity 185

7.2 Collaboration in geographical space 188

7.3 Collaboration in technological space 191

7.4 In a search of the gatekeepers 198

7.5 Conclusions 213

CHAPTER 8 STAR SCIENTISTS IN CANADIAN BIOTECHNOLOGY NETWORK 215

8.1 Identification of the star scientists in Canadian biotechnology 216

8.2 The positions of the star and QQ-star scientists in the network 223
8.2.1 Hypotheses 223
8.2.2 Methodology and results 226
8.2.3 Are stars and QQ-stars also gatekeepers? 230

8.3 Conclusions 232

9 CONCLUSIONS AND RECOMMENDATIONS 234

9.1 Concluding remarks 234

9.2 Recommendations for future research 240

REFERENCES 242

APPENDICES 277

LIST OF FIGURES

Figure 4-1: Shares of patents assigned to the various entities in each cluster (the number in parentheses shows the total number of patents in the cluster) 123

Figure 4-2: Proportions of patents by biotechnology field as granted to assignees in each category 134

Figure 4-3: Biotechnology field composition of patents in each cluster 135

Figure 4-4: Collaboration pattern of Canadian biotechnology institutions as evidenced by the patent co-assignment 137

Figure 5-1: Collaboration pattern in Canadian biotechnology 142

Figure 5-2: Collaboration pattern of the Canadian biotechnology clusters 143

Figure 5-3: Collaborations between Canadian and foreign inventors grouped by continents 144

Figure 5-4: Collaborations among Canadian inventors grouped by clusters 148

Figure 5-5: Proportions of all out-of-cluster collaborations (including both international and inter-cluster cooperation) based on the distance between the collaborators 153

Figure 6-1: Patents of Canadian biotechnology nanotechnology inventors by the year of granting 157

Figure 6-2: Patents and inventors in each biotechnology cluster based on the location of patent`s assignees and the residences of inventors 158

Figure 6-3: Patents and inventors in each nanotechnology cluster based on the location of patent`s assignees and the residences of inventors 158

Figure 7-1: Component C1 with all created C-C groups 206

Figure 7-2: Numbers and relative proportions of inventors in the clusters categorized according to their importance as intermediaries..211

Figure 8-1: Distribution of the number of patents authored or co-authored by each inventor...217

LIST OF TABLES

Table 1-1: Examples of definitions drawn from the literature by Martin and Sunley (2003) 8

Table 2-1: Comparison of the number of nanotechnology patents found by other authors 71

Table 2-2: Nanotechnology experts' answers to the question "Is this a nanotechnology related patent?" 72

Table 2-3: The final search strategy based on Porter *et al.* (2006) 75

Table 3-1: Results summary 80

Table 3-2: Results for the Porter externalities found in conjunction with Marshall and Jacobs positive results 81

Table 3-3: Number of indicators (independent variables) of Marshall externalities 83

Table 3-4: Number of indicators (independent variables) of Jacobs externalities 88

Table 3-5: Performance indicators by category of independent variable and industry classification level 90

Table 3-6: Number of positive results (dependent variables) per performance indicator 91

Table 3-7: Performance indicators per level of study by category of independent variable and industry classification level 98

Table 3-8: Number of positive results (dependent variables) per performance indicator and the level of study and 98

Table 3-9: Number of positive results (dependent variables) by industry sector type... 100

Table 3-10: Number of positive results per industry class by category of independent variable 105

Table 3-11: Number of positive results per geographical unit by category of independent variable and industry classification level 107

Table 3-12: Number of positive results (dependent variables) per geographical unit ... 107

Table 3-13: Number of positive results per country by category of independent variable and industry classification level 109

Table 3-14: Number of positive results (dependent variables) per geographical region (country) 110

Table 4-1: Summary of the results for biotechnology clusters 119

Table 4-2: Patents by category of assignees 121

Table 4-3: IP ownership and the distribution of the net revenue at the most prolific universities patenting in biotechnology in Canada 125

Table 4-4: Results regarding assignees as counted per biotechnology cluster 127

Table 4-5: Assignees with Canadian residence with 20 or more patents filed with the USPTO 129

Table 4-6: The number of the biotechnology patents and firms by the biotech field 133

Table 4-7: Patents by country of the assignees' residences 136

Table 4-8: Summary of information on the seven most important clusters 138

Table 5-1: Number of collaborations among the inventors in Canadian biotechnology clusters and in the US regions (slightly modified US Census Regions) 146

Table 5-2: Number of collaborations among the biotechnology clusters in Canada.....149

Table 6-1: Summary of the results for biotechnology and nanotechnology clusters.....160

Table 6-2: Statistics regarding collaborators or collaborations for each cluster............164

Table 6-3: Collaboration characteristics in biotechnology and nanotechnology cluster-based subnetworks.....166

Table 6-4: Fragmentation in biotechnology and nanotechnology cluster-based subnetworks.....170

Table 6-5: Structural cohesion and cliquishness in biotechnology and nanotechnology cluster-based subnetworks.....172

Table 6-6: Centrality and centralization in biotechnology and nanotechnology cluster-based subnetworks.....175

Table 6-7: Geodesic distances in biotechnology and nanotechnology cluster-based subnetworks.....178

Table 7-1: Structural properties of the cluster-based subnetworks (this is a summary of the results obtained in Chapter 6).....190

Table 7-2: Main characteristics and composition of the 30 largest components in the Canadian biotechnology innovation network.....193

Table 7-3: Structural properties of the component-based subnetworks.....195

Table 7-4: Inter-lines analysis for all intermediaries.....200

Table 7-5: Inter-line analysis for non-redundant intermediaries only.....201

Table 7-6: First 25 non-redundant intermediaries with highest inter-clines count showing values of all importance indices and other network properties.....204

Table 7-7: Average values of the indices of importance for the gatekeepers in each cluster 210

Table 8-1: Positions of all prominent inventors according to three different measures 219

Table 8-2: Prominent inventors by cluster 223

Table 8-3: The examined variables for each hypothesis 228

Table 8-4: Pearson correlation coefficients 228

Table 8-5: The results for all the hypotheses 230

Table 8-6: Overlapping of stars, QQ-stars and highly cited scientists with gatekeepers 232

LIST OF APPENDICES

APPENDIX A (Chapter 2): Example of questionnaire answered by a nanotechnology expert 278

APPENDIX B (Chapter 3): Summary of the studies providing empirical evidence.. 288

APPENDIX C (Chapter 3): Industries with positive results for Marshall externalities 305

APPENDIX D (Chapter 3): Industries with positive results for Jacobs externalities 306

APPENDIX E (Chapter 6): Biotechnology- Single factor ANOVA for the differences in population means 307

APPENDIX F (Chapter 6): Nanotechnology- Single factor ANOVA for the differences in population means 308

INTRODUCTION

In the last decade there has been a widespread resurgence of interest in the economics of industrial location and particularly in the issue of geographical clusters. Following successful cases in the United States (e.g. Silicon Valley) as well as Europe (e.g. Baden-Wurttemberg), governments of the industrialized countries have launched many programs with the aim of supporting regional innovation policies. To encourage innovative activities and promote competition the government's Innovation Strategy for Canada has decided to create at least ten internationally renowned technology clusters by 2010.

Knowledge flows are recognized to be a key explanatory factor for the geographical clustering of innovative firms and a critical element of the contribution of clusters to regional economic growth. The geographical and cognitive distance between the various academic or industrial inventors, networking structures of their collaboration activities and their own network positions all have a profound effect on the diffusion of knowledge through the network, and consequently on the performance of the firms within clusters in terms of their propensity to innovate. Therefore, in order to understand the key elements that support the growth of industrial clusters in Canada, a deeper understanding of knowledge diffusion in clusters and in the innovation networks is necessary. This thesis aims to shed light exactly on these issues.

The research in this thesis is organized around four main themes. The first theme concerns the debate on the economic development, growth and innovative performance of firms in the specialized versus diversified clusters or regions. Whether it is diversity or specialisation of economic activities which better promotes economic growth and innovation has been the subject of a heated debate in the economic literature which has yet to reach conclusive results. The findings of the investigation regarding the two concepts may play an important role in the design of a regional development strategy, since by turning the clusters either more diverse or more specialized a more favourable environment for the growth and innovative performance of the firms might be achieved. This thesis provides a census of the papers that have tried to contribute to the urbanisation versus localisation debate. The aim is not to try to determine which one of the two concepts better promotes innovation and economic development, but to investigate why it is that the literature still remains relatively inconclusive. The thesis therefore attempts to find the similarities and differences between the various studies in order to draw conclusions on the question.

The second theme of the thesis concerns the creation of innovation in Canadian biotechnology clusters. Canada has a small population dispersed over a large geographical area and its private sector is dominated by small-sized and medium-sized companies. As a consequence, research and development has to concentrate in geographical agglomerations and clusters in order to contribute to an efficient innovation system. The biotechnology field in particular should presumably benefit from the types of knowledge spillovers and information exchanges that are facilitated by spatial clustering. Growth and continued health of Canadian biotechnology clusters are among others dependent upon the presence of major attractors such as research universities and governmental laboratories active in biotechnology, innovative propensity of the local scientists, formation of alliances and active cooperation among biotechnology firms, composition of biotechnology fields in the cluster and presence of the largely innovative biotechnology firms. The thesis intends to address most of the above factors. It aims at understanding the creation of innovation in Canadian biotechnology clusters and its main objective is to identify, analyze and describe Canadian biotechnology clusters based on the characterization of the quality and quantity of their innovative outputs, the nature of biotechnology activities which are carried out in these clusters, the characteristics of the patent-owning entities and their propensity to collaborate.

The innovation networks and collaboration are the third major topic of this work. Innovation networks are important highways of information and knowledge travelling among various inventors or companies. Two relevant concepts are considered here – geographical and cognitive proximity. Geographical proximity facilitates knowledge sharing, since knowledge does not spill over large distances and the inventors collocated in the cluster can benefit from the local knowledge which is not available to the inventors outside the clusters. The second concept of knowledge creation and diffusion emphasizes the role of cognitive proximity, based on which it is not geographic proximity which causes tacit knowledge to spill over between firms, but it is social connectedness of people in the network. Knowledge circulates and flows through the innovation networks among the inventors who are not necessarily placed in the same location. This thesis brings together the findings on both the importance of geographical proximity and the significance of the structure of the innovation networks. The transmission of knowledge through collaboration inside and outside Canadian biotechnology and nanotechnology clusters is examined and the role of the structure of the collaboration networks in the creation and diffusion of innovation is investigated.

Finally, the fourth theme deals with the prominent inventors in Canadian biotechnology clusters and their roles in the innovation networks. Two kinds of prominent inventors are considered – gatekeepers and star scientists. In this thesis it is suggested that both geographical and cognitive dimensions nurture the growth of the cluster and promote innovation through a dynamic interaction of the actors localized in clusters who absorb external knowledge through the local and non-local networks. Therefore the inventors who are well connected both inside and outside the clusters are needed in order to bring the new knowledge to the cluster. These inventors are called gatekeepers and they are the points of interaction between the geographical and technological (network) collaboration spaces. Star scientists, on the other hand, may not be necessarily well connected in the network, but they have to be extraordinarily highly prolific in their research and inventive productivity. These top scientific individuals are recognized as a key driving force for the growth and innovation in biotechnology. The presence of star scientists often even explains the timing, location and the success of new biotechnology firms. This thesis searches for both the gatekeepers and the star scientists in the Canadian biotechnology clusters and also investigates their geographical and network positions.

The thesis is organised as follows: Chapter 1 introduces the main topic of the thesis by reviewing the relevant research work from the literature. Chapter 2 first presents the research questions, explains concrete research objectives and then describes the general organization of the thesis. It also discusses the methodology and the data used in this work. Chapter 3 concerns the debate on the economic performance and growth of firms in the specialized versus diversified clusters or regions. Here, the main aim is to provide taxonomy of scientific articles and to investigate the reasons behind the inconsistency of the findings. The second objective of the thesis pertains to the creation of innovation in Canadian biotechnology clusters. The goal of the Chapter 4 is thus to identify, analyze and characterize these clusters, with a special focus on the intellectual property politics in Canadian institutions. The innovation networks and collaboration are third major theme of this work. In Chapter 5, the thesis examines the geographical aspects of the collaboration and the impact of geographical proximity on the selection of the collaboration partners in Canadian biotechnology and nanotechnology. Afterwards, in Chapter 6, the characteristics of the biotechnology and nanotechnology collaboration networks are compared and their role in the efficient knowledge diffusion highlighted. Last two objectives are related to the prominent inventors in Canadian biotechnology clusters and their roles in the innovation networks. Chapter 7 presents innovative methodologies for the identification of the knowledge gatekeepers in the collaboration network. The last aim concerns the star scientists. Chapter 8 searches for different ways for their identification and studies their geographical and network positions. Finally, Chapter 9 concludes, describes the contributions to the advancement of knowledge and proposes several avenues for future research.

The main approach of the thesis consists in the creative exploitation of the large amounts of information extracted from the patent database of the United States Patent and Trademark Office (USPTO). The information is used to describe the creation of biotechnology and nanotechnology innovation in Canada and to build the innovation networks based on the patent co-inventorship links. The methods of social network analysis are used to create, characterize and evaluate these innovation networks.

The sectors selected for the analysis are biotechnology and nanotechnology. They belong among Canada's most dynamic sectors and provide a significant contribution to science advancement and innovation, thousands of jobs, as well as large exports. This makes the subject of this research very important for Canada.

CHAPTER 1

LITERATURE REVIEW

The objective of this survey is to introduce the main topic of the thesis by reviewing the relevant research work from the literature. It is divided into three main sections, focusing on different research area of the principal research theme:

The first section introduces the concept of clusters, provides a basic rationale behind the clustering phenomenon and offers a survey of the published findings regarding the performance of the firm in a cluster.

In the second section, the topic of localized knowledge spillovers is presented in detail. The starting point here is the review of the literature pertaining to the knowledge production function and to the most researched and relevant aspect of knowledge spillovers – their localization effect. This is followed by a discussion on public science (public research), as a main source for the knowledge spillovers. A review of knowledge properties is presented afterwards in order to help better understand the various mechanisms leading to the generation of knowledge spillovers. The section also provides a survey of the critical papers reassessing some of the main ideas regarding knowledge spillovers theory and their localization effect. Finally, the topic of diversification and specialization externalities is introduced.

The third section is devoted to the publications pertaining to innovation networks. The concept of collective invention is first discussed, then the literature studying the structure of the networks of innovators and inter-firm collaborative networks is reviewed. The survey of research concerning the key individuals in the network - the brokers and gatekeepers – concludes this section.

1.1 Clusters

A cluster is defined by Porter (1998) as a geographic concentration of interconnected companies, specialised suppliers, service providers, firms in related industries, and associated institutions (for example, universities, standards agencies, and trade associations) in particular fields that compete but also co-operate. Nevertheless, the use of the term "cluster" is not completely standardized. The above definition is vague and flexible in terms of geographical scale and internal socio-economic dynamics, which gives rise to a wide range of interpretations found in the literature. Martin and Sunley (2003) claim that "clusters have no essential self-defining boundaries, whether in terms of inter-sectoral or inter-firm linkages, information networks, or geographical reach. The notion is so generic that it is used as a sort of cover term to a whole assortment of types and degrees of industrialized localization."

Table 1-1 shows some instances of distinct definitions which Martin and Sunley (2003) encountered in the literature. For example, the definition of the cluster may be based solely on the geographical dimension (Definition 10). Much more common however is to define the cluster both by the proximate location of firms and by the similar or related type of industrial fields (Definitions 1, 2, 5 and 6). Many definitions stress the importance of linking, relationships or interconnections among the companies within the cluster (Definitions 1, 3, 4, 7, 8, and 9). Out of these definitions, some (Definitions 4, 7 and 8) even do not require the condition of a geographical proximity, since it is the high degree of collaboration or interdependence which seems to be playing much more important role. Martin and Sunley (2003) suggest that the ambiguity in the definitions has allowed different analysts to use the idea cluster in different ways to suit their own purposes.

Table 1-1: Examples of definitions drawn from the literature by Martin and Sunley (2003)

Definition #	***Definition***
1	"A cluster is a geographically proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and complementarities".
2	"The more general concept of 'cluster' suggests something looser: a tendency for firms in similar types of business to locate close together, though without having a particularly important presence in an area."
3	"A cluster is very simply used to represent concentrations of firms that are able to produce synergy because of their geographical proximity and interdependence, even though their scale of employment may not be pronounced or prominent."
4	"Economic clusters are not just related and supporting industries and institutions, but rather related and supporting institutions that are more competitive by virtue of their relationships."
5	"Clusters are here defined as groups of firms within one industry based in one geographical area."
6	"A cluster means a large group of firms in related industries at a particular location".
7	"We define an innovative cluster as a large number of interconnected industrial and/or service companies having a high degree of collaboration, typically through a supply chain, and operating under the same market conditions."
8	"Clusters can be characterised as networks of producers of strongly interdependent firms (including specialised suppliers) linked to each other in a value-adding production chain."
9	"The popular term cluster is most closely related to this local or regional dimension of networks ... Most definitions share the notion of clusters as localised networks of specialised organisations, whose production processes are closely linked through the exchange of goods, services and/or knowledge."
10	"A regional cluster is an industrial cluster in which member firms are in close proximity to each other."

1.1.1 Clustering benefits and costs

The original theories about the emergence of clusters come from Marshall (1890); however, the basic idea behind the clustering mechanism was explained by Krugman (1991). He has defined three sources of geographical concentration of industries, which stimulate entry into regions that have previously accumulated a large number of firms, as labour market pooling, availability of intermediate inputs and knowledge spillovers. These sources represent the supply-side benefits of clustering, because they refer to the production process of a firm. The existence of a pool of labour specialized in particular technical or scientific knowledge and skills can significantly lower the company's search and transaction costs in recruiting within cluster. A cluster attracts companies, which in turn create more specialized labour in that cluster. As a consequence, attracting talented people from other locations to such clusters becomes even easier and their search and recruitment cheaper. In some industries these economies of labour pooling may create a decisive competitive advantage (Porter, 1998). A specialized supplier base, which provides a company with an efficient way to obtain many important specialized inputs, is usually present in every developed cluster. Specialized inputs are any inputs of equipment, research tools and related technologies that need to be tailored and developed for a particular market (Prevezer, 1997). Locating near a pool of the specialized inputs allows a firm to obtain a much greater variety of those inputs and at lower costs. Knowledge spillovers are the third and the most discussed of the supply side benefits. Knowledge spillovers can be defined as "positive externalities of scientific discoveries on the productivity of firms which neither made the discovery themselves nor licensed its use from the holder of intellectual property rights" (Zucker *et al.*, 1998a). Knowledge and information flow more easily between firms located in a cluster than over long distances (Audretsch and Feldman, 1996) and firms therefore locate close to the sources of spillovers that are essential to their activity. Other important agglomeration benefits of the location in a cluster are sharing a physical infrastructure and communication technologies. Also, decreased transportation costs of inputs needed by the firms to produce their own product or lower transportation costs to the consumer markets are additional positive effects, which may help explain the existence of clusters (Porter 1998).

Baptista and Swann (1998), who surveyed the factors that enhance and cause clusters, distinguish four benefits at the demand side. These are strong demand, market share gain, lower search costs and customer's feedback. Clusters may arise at places with strong local demand, which is often deriving from the related industries present in the cluster. Local demand and a great market may attract companies from the same, similar or complementary sectors. Firms within one sector may gain market share from locating close to other established firms. By moving closer to their rivals, companies may capture the share of the market that is serviced by the competitors. The existence of a cluster decreases search costs for customers, because clusters of firms allow customers to assess and compare firms and their products more easily. Certain specialized companies with differentiated goods may find it very advantageous to move to the cluster in which they could be more easily spotted and discovered by the customers. Firms located near customer markets can also exploit information flows of important customers, who could become a good source of innovation ideas. For example, the firms may decide to provide additional customer services according to the customers' wishes.

Since this set of advantages is relatively immobile, firms choose to move from other locations to existing clusters in order to capture the benefits. This creates further positive feedback for other companies and leads to the growth of clusters through a self-reinforcing process. The benefits from clustering are, however, limited by the negative effects, which are increased competition, the congestion effects or technological discontinuities (Baptista and Swann, 1998). More intense competition between firms within a sector in producing the same product will drive down pricing power of the companies, which will lead to the reduction of the firm's profits, sales, etc. Congestion effects can cause increased prices of housing, wages or land rents, and consequently increased production costs and lower profits. Overgrown clusters may generate other negative externalities, for example due to the increased pollution or overcrowding. The technological discontinuities, which occur when new technologies appear and the old ones are taken over, may lead to the decline and even the death of the cluster.

1.1.2 Performance of the firm in cluster

There is a great amount of work focusing on the dynamics of cluster generation. Many models that study the influence of the strength of the industrial cluster on the performance of the company located in a cluster were developed and presented in the econometric literature. There are three signals of a successful cluster according to Porter (1998): rapid firm growth, new firm entry and innovation. The following studies model the firm's growth, entry or innovative activity as a function of the strength of the cluster in which it is located and evaluate the effects of clustering.

The growth of incumbent firms in a cluster was examined by Swann and Prevezer (1996) using data from two industries: computing and biotechnology. For both industries they found that company growth is promoted by industry strengths in its own sector, while the role of the strength in other sectors or in the science base was found negligible. Baptista and Swann (1998) continued this research focusing on the same two industries and confirmed the previous results regarding faster firm growth in its own industry clusters. Furthermore, they found that firms located in clusters that were strong in other industries did not grow faster and sometimes might even grow slower. This is suggested to be an indication of congestion effects that outweigh any possible benefits coming from diversification within clusters. This agrees with the results of Beaudry (2001) who studied the relevance of the conclusions from these studies in the context of the aerospace industry in UK. She confirms both the positive impact of own-sector clustering and the negative impact of other-sector clustering on firms' growth rate in most industrial sectors. In a study on clustering in the US and UK computer industries Baptista and Swann (1999) again validated the previous findings. They also added that firms in generally strong clusters tend to grow faster.

Swann and Prevezer (1996) also studied the firm's **entry** to the cluster and found distinct results for computing and biotechnology industries. They discovered some important cross-sectoral effects that promoted entry to the computing industry, while in biotechnology these cross-sectoral effects were more limited. Moreover, it is argued that new firms in biotechnology are strongly attracted to the presence of a strong science base at the location, which was also confirmed by Prevezer (1997). Her other interesting finding was that new companies were attracted by the entry of other new companies, except in their own industrial sectors. Prevezer argued that the prospect of competition at a location within its own sector acts as a deterrent to a new firm from setting up at that location. In addition, she observed that clusters of biotechnology firms develop only in particular sectors of the industry. This is consistent with the results of Baptista and Swann (1999) who observed that in both the US and UK computer industries, new companies are attracted by industry strength in particular sub-sectors in a particular region, and also with the findings of Beaudry's (2001) research on clustering in the UK aerospace industry, where she observed that some sub-sectors of the industry attract new entry while others are only attracted. The results from both computing and biotechnology industries studied by Baptista and Swann (1998) confirm that the strongest attraction effects are across sub-sectors of each industry. Finally, Baptista and Swann (1999) also found that the clusters that are more likely to attract new entrants are usually the strong ones. By entry of new firms the cluster thus becomes even stronger and attracts more other firms. It is argued that this cluster self-reinforcing effect could start out of the emergence of one strong firm. Wolfe and Gertler (2004) emphasize the importance of an anchor firm for the cluster and give practical examples when entire clusters developed out of the formation of one or two critical firms. The anchor firms attract both allies and rivals to the region to monitor the activities of the dominant firm.

The research evaluating the effect of clustering on the innovation rate has shown similar results as the previously discussed studies on the firm's growth in clusters. Both the positive effect of own-sector clustering and the negative impact of other-sector clustering on the number of generated innovations were observed by Baptista and Swann (1998) in biotechnology and computing, by Beaudry (2001) in aerospace industry and recently also by Beaudry and Breschi (2003). Moreover, the latter article emphasized that clustering in itself will not necessarily lead to higher innovative performance. The authors observed that the probability of innovation for a firm is much higher if it is located in a region with a large accumulated stock of knowledge. The cumulative nature of the innovative activity has been suggested also by Arthur (1990) who claimed that a key aspect of the effect of clusters on a firm's innovative activities is the accumulated stock of knowledge in a particular area. This is also in agreement with Lamoreaux and

Sokoloff (1997) who also observed that inventive activity will tend to concentrate in locations where invention rates had long been high.

The observation that innovative activities are strongly geographically agglomerated has thus led many researchers to investigate the likely causes of this phenomenon. The following section discusses localized knowledge spillovers as a key explanatory factor of local clustering.

1.2 Localized knowledge spillovers

Authors of econometric studies of the geography of innovation have frequently claimed that localized knowledge spillovers are a key explanatory factor for the geographical concentration of innovative activity (Dahl and Pedersen, 2004). Localized knowledge spillovers can be defined as knowledge externalities bounded in space, which allow companies operating nearby key knowledge sources to introduce innovations at a faster rate than rival firms located elsewhere (Bresch and Lissoni, 2001b). Knowledge developed in a cluster or industrial district flows more easily within it, but more slowly outside and across its borders. And since geographic proximity reduces the cost of accessing and absorbing knowledge spillovers, the innovative activity will tend to geographically concentrate close to agglomerations of the mentioned infrastructure in order to benefit from spillovers (Bresch and Lissoni, 2001a).

1.2.1 The concept of the knowledge production function

A fundamental issue which remains unresolved in the economics of technology is the identification and measurement of knowledge spillovers coming from research activity, specifically the extent to which a firm is able to exploit economically the investment in research made by other party, as university, public research institution or another company. The traditional way of providing evidence of the existence of the knowledge spillovers has used the knowledge production function.

The model of the knowledge production function, formalized Zvi Griliches (1979), simply states that innovative output is a function of innovative inputs. The most important source of new knowledge is considered to be R&D, other factors are human capital - a skilled labour force, scientists and engineers. The degree of innovative activity is therefore a function of the amount of R&D expenditures and human capital inputs.

The unit of observation for estimating the model of the knowledge production function could be at the level of countries, industries, clusters or enterprises. However, empirical estimation of the model of the knowledge production function was found to be stronger at broader levels of aggregation such as countries or industries. If the unit of observation is countries, the relationship between R&D and patents is very strong (the most innovative countries as Japan, USA or Germany have also high investments in R&D). Also for the industry as an observation unit, the link is very strong: the most innovative industries, computers, instruments and pharmaceuticals, are also R&D intensive. However, if tested for the firm as an observation unit, the link between innovative input and output becomes only weakly positive, non-existent or even negative (Audretsch, 1998). Formal R&D is usually undertaken by the large and established corporations, but some studies (Acs and Audretsch, 1990; Audretsch, 1995; Scherer, 1991) have documented that small and new firms that do not carry out much of the formal R&D themselves still generate a substantial innovative activity, especially in newly emerging industries such as biotechnology and computer software.

An explanation for the disproportionate share of new product innovations of small firms (given their low R&D expenditures) has recently emerged in the economic literature. It is suggested that it is from other, third-party firms or research institutions, such as universities or governmental laboratories conducting R&D, where new knowledge may spill over and innovative firms with little or no R&D may appropriate the knowledge inputs. The following section briefly summarizes important findings which suggest that investments in R&D by private corporations and universities "spill over" for third party firms to exploit.

1.2.2 Localization effects of knowledge spillovers

Several researchers provided empirical evidence that location and proximity are an important factor in exploiting knowledge spillovers. Jaffe (1989) was the first one who found a sign of the existence of localized technological spillovers from academic institutions into local enterprises. He modified the knowledge production function introduced by Griliches (1979) and shifted the model of production function from the unit of observation of a firm to that of a geographical unit. He showed that the number of patents of each US state for each technological area is a positive function of the R&D performed by local universities. The knowledge production function used together with Jaffe's geographic coincidence index for analysis of local spillovers then became a common tool for the study of the localization effects of knowledge spillovers and effects of local university research on the innovative activity of the companies. The following studies all use this framework. Acs *et al.* (1992) carried out a similar research as Jaffe (1989) focusing on electronics and mechanics industry sectors. They introduced the measure of innovation counts using US Small Business Database (SBA) and proposed it to be a better indicator of innovative output than previously used patents (For discussion on patent counts as innovative indicator see 1.4.1, for innovation counts see 1.4.2) They also confirmed that university research has a strong effect on patenting of enterprises. The findings of Acs *et al.* (1994) suggest that the innovative output of all firms rises along with an increase in the amount of R&D inputs, both in private corporations as well as in university laboratories. However, they observed that knowledge spillovers are not homogenous across firms and proposed two different knowledge production functions, one for large firms and one for small ones. Audretsch and Feldman (1996) changed the focus from the product dimension to a geographic or spatial dimension and showed that the R&D intensity of the industry is positively influenced by the geographical concentration of the innovation output. They also concluded that knowledge externalities are more prevalent in industries where new economic knowledge plays a greater role.

Several authors afterwards confirmed that the innovative activity has a propensity to cluster spatially and suggested the existence of the knowledge spillovers, still using the knowledge production function concept. Anselin *et al.* (1997) introduced the use of metropolitan statistical areas (MSA) in the framework of knowledge production function. They refined Jaffe's geographic coincidence index for analysis of local spillovers and proposed the research concept that provides an evidence of the effects of localized knowledge flows on regional innovation. They found the indication of geographic spillovers from university research to innovations and indirectly to industry research. The authors observed that spillovers from university research extended over a range of 50 miles from innovating MSA, but not with respect to private R&D. Acs *et al.* (2002) extended their previous work (Anselin *et al.*, 1997) and confirmed their results about the existence of the localized knowledge spillovers. The central finding of their paper was that the two measures of technological change (patents and innovations) produce very similar results in regression models of regional spillover activity. Similar method is used by Fisher and Varga (2003) to investigate the effect of university research on patenting in Austrian political districts. Their results provide evidence of mediated knowledge spillovers from university research to the production of regional knowledge. Spillovers cross political districts and clearly decrease in intensity with distance. Kelly and Hageman (1999) showed a strong spatial clustering of the patenting activity using different methodology. They developed a quality ladder model and found that innovation exhibits strong spatial clustering independently of the distribution of employment. They concluded that the innovative performance of the state is greatly influenced by the existence of knowledge spillovers.

Another stream of research on knowledge spillovers focused on tracking the knowledge flows (usually from academic research into corporate R&D) with the use of the patent citations as a representation for knowledge spillovers (for more information regarding the use of patent citations and their use as indicators of knowledge spillovers see 1.4.1). The following studies show evidence of a localization effect of patent citations, implying that knowledge diffusion is geographically localized. Jaffe *et al.* (1993) found that patent citations tend to occur more frequently within the state in which they were patented than outside that state, which means that innovative firms are more likely to quote research from a co-localized university that conducts relevant research, than from similar universities located elsewhere. However, they also found evidence that geographic localization fades over time. Jaffe and Trajtenberg (1996) developed a model of the process generating subsequent citations to a patent to represent knowledge diffusion. The results indicate that knowledge diffusion is geographically localized. The research of Almeida and Kogut (1997) examined the innovative ability of small firms and the geographic characteristics of spillovers using the patent citations. Their findings revealed that small firms are tied into regional knowledge networks to a greater extent than large firms, and that knowledge spillovers are highly localized. Maurseth and Verspagen (1999a) used patent citations to study the knowledge flows between the regions and confirm that the number of citations rapidly decreases with distance. Maurseth and Verspagen (1999b) found an evidence of national barriers to citations. They observed that citations occur much more frequently between regions within national states than to regions belonging to other countries, which was also confirmed by Jaffe and Trajtenberg (1996). Verspagen and Schoenmakers (2000) extended the work of Maurseth and Verspagen (1999a) and tested for the proximity effect by measuring geographical distance at the level of firms, using data on the location of inventive activities of the firms. They again validated the geographic proximity effect of patent citations. Further research was carried out by Jaffe *et al.* (2000), who surveyed a number of inventors. They also found clear evidence of a localization effect of patent citations, meaning that knowledge diffusion is geographically localized.

Previous research has shown that the localization effects of knowledge spillovers vary across industries. It is argued that the importance of tacit knowledge in the industry is one of the factors that determine the industry concentration. Audretsch and Feldman (1996) found that a key determinant of the extent to which location of production is geographically concentrated is the relative importance of new economic knowledge in the industry. They concluded that in industries where new knowledge plays a crucial role, innovative activity tends to cluster in locations where key knowledge inputs are available.

Prevenzer (1997) carried out a study to identify the forces of attraction to new companies to a cluster in biotechnology sector, which is an industry based almost exclusively on new knowledge. They found that unlike the companies in other industries the biotechnical firms tend to cluster together in only several locations. The main agent of attraction to new firms to enter the biotechnology industry is identified as the presence of a strong science base at that location. Audretsch and Stephan (1996) support these findings when they examine the geographic relationships of scientists working with biotechnology firms. They suggest that specific role played by the scientist shapes the importance of geographic proximity in the link between firm and the scientist. When the scientist's role includes a transfer of tacit knowledge, local proximity is much more important than if the knowledge is codified.

1.2.3 Knowledge flows from public science (public research)

Public science (or public research) in this thesis is understood to be the knowledge that originates from universities, research institutions, government laboratories, etc. It is widely accepted that public sector research makes a significant contribution to growth by supplying basic non-market oriented scientific knowledge that the private sector has weak incentives to produce. Recent research (McMillan *et al.*, 2000) has shown that the overall US industrial base relies heavily on external sources of knowledge centers, on public science. Narin *et al.* (1997) found out that during 1993-

1994, 73% of the scientific papers by US industrial patents were from public science sources, while only 27% were authored by industrial scientists. The role of public science is crucial specifically in certain industries. For example, Zucker and Brewer (1994) claimed that science was in fact an external stimulus to the founding of the biotechnology industry. Biotechnology originated from a series of scientific discoveries and the science base has remained a critical source of innovation in this field (Prevezer, 1997).

In recent years there has been a great deal of interest in the process by which firms benefit from externally performed research and development, and the extent and importance of such spillovers. In the following sections the aim is to present a review of the literature which concerns the knowledge flows from public science into the private sphere.

1.2.3.1 Science and technology environment

Dasgupta and David (1994) described the differences between the social organization of the worlds of science and technology: Science is characterized by publication, supported by a priority-based reward system and exists mainly in research universities. This is a contrast with the world of technology in which ideas are produced for economic objectives and encoded in patents and other modes of protection to facilitate appropriability. Balconi *et al.* (2004) emphasizes the difference in openness of the two environments. Within the world of technology, the results, instruments and methods are shared with other researchers, but not outside organizational boundaries. Communication with rival companies is monitored and restricted and codification efforts are delayed as long as possible. By contrast, each group of academic scientists belongs to a wide community of researchers of the same field and contributes to expanding, codifying and securing the reliability of scientific knowledge. Murray (2002) reached similar conclusions when her research showed that the scientific and technical networks are quite distinct. They differ in several aspects including size, workforce, institutions and the nature of collaboration.

Despite the differences in the world of science and world of technology, scientific and technological ideas in fact co-evolve. Murray (2002) analyzed the dynamics of such co-evolution and discovered that the co-mingling is carried out mainly through firm founding, licensing, consulting and advising, and not through co-publishing or citations, as was predicted. Only few key scientists publish across industry-academic boundaries and firms in fact do not participate in science. Zucker *et al.* (1998a) confirm that especially among scientists it is commonly thought that the very best scientists are unlikely to be involved with the firms or to patent their discoveries. Dasgupta and David (1994) also point out that knowledge transfers from university-based open science to commercial science are quite inefficient. Part of this inefficiency is a consequence of the constant friction between academic institutions who desire publication and the establishment of priority, and corporate research sponsors who wish to defer disclosure until appropriate mechanisms such as patent can be employed to protect the future economic returns of an innovation.

1.2.3.2 Academic research

Industrial innovation relies heavily on sources of basic scientific knowledge coming from university research. In his study based on data obtained from 76 firms from 7 industries, Mansfield (1995) found that about 11% of their new products and about 9% of the new processes could not have been developed without the findings of recent academic research. In the absence of the academic research there would be substantial delay and much higher costs, which would often make the new product development economically undesirable. Mansfield (1998) continued his research with a focus on the change in trend over time and reports an increase in the percentage of new products and processes based on academic research in 1986-1994 relative to 1975-1985. Research by Acs *et al.* (1992, 1994) and others also confirms that technological change in important segments of the economy has been based significantly on academic research.

Geographical proximity to universities gives direct access to individuals that can efficiently turn information into usable knowledge, making commercial control over a technology easier and faster. New knowledge and technological-based firms have therefore a high propensity to locate close to universities. It is presumed that they do so in order to access knowledge spillovers coming from the academic institution. University spillovers are defined by Harris (2001) as externalities towards firms, for which the university is the source of the spillover but is not fully compensated.

Jaffe (1989) constructed a model to identify the contribution of university research to creating innovation. His statistical results provide evidence that corporate patenting at the state level depends on university research spending. Not only patent activity increases in the presence of high private corporate expenditures on R&D, but also as a result of research expenditures undertaken by universities within the state. Liebeskind *et al.* (1996) explored the situation in the biotechnology companies. They concluded that companies who engaged in joint research and publishing with academic institutions were more effective at sourcing new scientific knowledge than those who did not have joint activities.

However, Mansfield (1991) was initially hesitant to acknowledge the importance of the local university for the corporate research. He surveyed industrial R&D employees about university research from which they benefited. He found that even though there was some tendency to cite local universities even if they were not the best in their field, they most often identified major research universities. Nevertheless, Mansfield (1995) later extended his research and identified more precisely the factors that determine how much the university research contributes to innovation in the companies. He found out that the extent of the contribution is related directly to the quality of the university faculty in the relevant department, to the size of its R&D

expenditures in relevant fields and to the proportion of the industry members located nearby.

The current research concerning knowledge flows from academia and university spillovers focuses on the propensity of a new firm to locate within a close proximity to the university. Audretsch and Lehmann (2005) and Audretsch *et al.* (2005) identified the factors that increase the attraction power of the universities for the new firms and their influence on the locational strategy of a firm. The empirical evidence provided by Audretsch and Lehmann (2005) suggests the number of firms located close to a university is positively influenced by the knowledge output of a university, which confirms the mentioned findings of Mansfield (1995). The authors also claimed that the universities located in the region with a high regional investment in knowledge tend to attract more technology startups. The results of Audretsch *et al.* (2005) show that the impact of university output on new firm location is sensitive to both the type of knowledge and the mechanism used to access that knowledge. They found that new firms do not have a high propensity to locate within close proximity to universities with a high research output in the natural sciences, while the propensity is much higher for the universities which focus on the research in social sciences. It is explained by the properties of knowledge in natural sciences, which is much more codified and therefore distance insensitive, whereas knowledge transmitted through published research in the social sciences is more tacit, leading new firms to locate closer to the university in order to access the knowledge spillover. The results were however opposite when they examined another spillover mechanism, which is human capital. New firms tend to locate more closely to universities with a large output of students in the natural sciences; but that does not hold for social sciences. The authors explain that this is caused by the fact that human capital in the natural sciences is more specific and less general than in the social sciences.

The intensity of the university spillovers flowing into the companies is not influenced only by the characteristics of the university and the research that is conducted there, but it also depends on the size of the firm which is the recipient of the spillovers from knowledge generated in the R&D centers of the universities. The findings of already mentioned research conducted by Acs *et al.* (1994) provide substantial empirical evidence that spillovers from university research laboratories are more important in producing innovative activity in small firms, whereas corporate R&D is a relatively more important source for generating innovations in large firms. This agrees with Link and Rees (1990) who reported that small new entrepreneurial firms tend to benefit more from university research spillovers than larger and established corporations.

1.2.3.3 Nature of university research as evidenced by patent data

There is a stream of the economics of innovation literature that focuses on the study of patenting, patents and patent citations. The many advantages of the use of patents to evaluate innovative activity and more detailed analysis of this method could be found in section 1.4. The following findings are the results of the works which studied the patent data as a manifestation of inventive activity in order to determine and analyze the nature of research and development in universities.

University patenting is in fact insignificant, compared with the patenting of companies and other institutions. Trajtenberg *et al.* (1997) stated that university patents account for a very small fraction of all patents, for example in 1990 it was only 1.2% of all patents granted in the U.S. that year. However, the R&D performed by academic institutions in the US constituted in the same year 11.4% of total R&D share expenditures. This suggests that the patenting activity of universities per dollar of R&D expenditures is very low, but is explained by the distinct nature of academic research (basic research) and incentives in academia (preference of scientific articles).

As for the evolution of the patenting over time, Henderson *et al.* (1998b) have studied the pattern of university patenting in the U.S. in the period of 1965-1988 and have shown the number of university patents increased. They suggest that this increase in university patenting probably reflects an increased rate of technology transfer to the private sector. At the same time, however, the steady growth in university patenting has been accompanied by a steady fall in the average quality of university patents, whose relative importance has declined. Before about 1985, university patents on average were much more highly cited than other patents, this difference, however, almost disappeared by the late 1980s. According to the findings of Hicks *et al.* (2001), since 1993 university patents are less frequently cited than US company patents, and since 1999 they are even less cited than an average patent. Henderson *et al.* (1998a) explained that the decline in relative quality of university patents has been probably driven by a reduction in the standards for patenting as incentives changed. Bayh-Dole Act passed in 1980 gave universities the right to retain the intellectual property rights to all the inventions, and their propensity to patent consequently increased. Instead of patenting only their most significant innovations universities have moved on to patent less significant research output as well.

University patents are highly concentrated in the hands of relatively few academic institutions: ten universities with highest number of patents received over 50% of all patents (Trajtenberg *et al.*, 1997). Henderson *et al.* (1998b) stated that the top 20 universities received about 70% of the total number of patents granted to academic institutions, and MIT alone accounts for 8% of these patents. In fact, according to Hicks *et al.* (2001), MIT is the largest producer of patents in Boston and Harvard the fifth largest, while in San Francisco the University of California is the second largest patentee and Stanford the ninth. Although university patents form a small percentage of total national patenting, universities dominate patenting in some of the most economically vibrant large cities.

In addition, the university patents are also concentrated in a relatively small number of fields. At least 25-30% of university patents belongs to patent classes related to biological and medical sciences, which commanded 45% of all academic R&D in 1980 (Trajtenberg *et al.*, 1997). This agrees with Hicks *et al.* (2001) who confirm that it is in health technologies where universities achieve their most significant patenting presence (with a 15% share of the combined patenting from universities, government and industry).

The nature of the research done at academic institutions is widely assumed to be more basic, while private institutions are usually engaged in more applied efforts. Trajtenberg *et al.* (1997) defined about 65% of the university research as basic research, 30% as applied research and just 5% as development (in 1992). The findings of Jaffe and Lerner (2001) are in general agreement with these numbers, since they claim that two thirds of the university research is basic research, while it is about 40% of all federal lab research. The prevalence of the basic research in universities was also confirmed by Trajtenberg *et al.* (1997) who suggested that university research is located closer to the origin of the innovation path. They found that compared to the inventions patented by universities, the corporate innovations rely on a higher number of preceding inventions which are of higher economical value. University research relies relatively more on scientific (non-patent) sources than corporate research. These findings imply the basicness of the university research. Until recently, the basic nature of the academic research could have been also confirmed by the fact that university inventions were more cited and thus used more for further applications. Jaffe and Trajtenberg (1996) carried out a study on the fertility of university patents and found that they were more highly cited than corporate and federal patents. However, as the latest results (Hicks *et al.*, 2001) suggest, this is no longer true.

1.2.3.4 Governmental research

A federal research institution is an institution, which is operated by, or receives most of its funding from, the federal government. Federal research institutions are an important part of the U.S. research infrastructure. In 1995 in the U.S. 41% of federal spending on R&D was performed by federal research institutions, while universities received only 21% of federal research expenditures (Jaffe *et al.*, 1998). Jaffe and Lerner

(2001) stated that in the period of 1955-97, only 24% of the total federally funded R&D took place in academic institutions, whereas the majority of the research activities were in fact performed in governmental laboratories. Therefore it seems surprising that the university research has been studied intensively, while there is not a lot of published literature dealing with governmental laboratories.

The nature of R&D of federal research institutions is mission-oriented, and therefore they usually have a lower propensity to patent. Jaffe *et al.* (1998) found that the governmental research institutions generate many fewer patents per dollar of R&D than the private sector. Moreover, the governmental inventions which get patented do not seem to be of a very high economical value. The findings of Jaffe and Trajtenberg (1996) in their study on the patent fertility suggest that in the U.S. the federal government patents are significantly less highly cited than corporate patents, which are less cited than the university patents. Nevertheless, the patents generated by federal research institutions are usually cited for a longer period of time.

Jaffe *et al.* (1998) examined the patenting behaviour of NASA and other federal agencies over the last several decades, together with the average impact of these patents. They found an evidence of increased patenting activity by these agencies in the last decade; however they did not find any evidence that the increase in federal patenting would be associated with the decline in the average impact of the federal patents. This is not analogous with the already discussed findings of Henderson *et al.* (1998b) which found an increase in patenting by universities since the early 1980s accompanied by a significant decline in the average impact of university patents. The findings of Jaffe *et al.* (1998) are supported by Jaffe and Lerner (2001), who investigated the commercialization of publicly funded research in the U.S. national laboratories. They conclude that the policy reforms in the US in 1980s had dramatic and positive effects on technology commercialization and caused patenting to increase sharply, but the overall increase in patenting of national laboratory institutions was not associated with an overall decline in quality, as is the case of universities.

1.2.3.5 Absorbing knowledge spillovers

Cohen and Levinthal noted in two articles (1989 and 1990) that firms which want to take advantage of research conducted outside their organizational boundaries may need to invest in "absorptive capacity", which is explained as a need to accumulate the knowledge, skills and organizational routines necessary to identify and utilize externally generated knowledge. Cohen and Levinthal define absorptive capacity as "the ability of a firm to recognize new information, assimilate it, and apply it to commercial ends". The authors suggested that since an innovation process in a company is comprised of both internal and external elements, the exploitation of basic scientific discoveries requires an organization to continuously learn from beyond its boundaries. Most of the studies describing the dependence of organizations on external knowledge to enhance their studies consequently focus on "absorptive capacity" model defined by Cohen and Levinthal (1990).

Henderson and Cockburn (1996) agreed with the absorptive capacity model of Cohen and Levinthal (1990) and explored the idea further. They suggested that it may be necessary not only to invest in basic research inside the firms, but also to hire the best possible research personnel, which they call "star scientists". They claimed that increasing the quality of the human capital in the firm will improve internal research productivity. However, the authors also showed that substantive difficulties in measuring the quality of human capital make it difficult to estimate this effect precisely. They proposed the idea to reward the researchers on the basis of their standing in the public rank hierarchy (Henderson and Cockburn, 1994). They argued that firms that are pro-publication in the sense that they promote researchers on the basis of their standing in the scientific community are significantly more productive than their rivals, all other things equal. They also claimed that this rewarding system is more efficient (cheaper), because it forces researchers to publish and stay in touch with the state of knowledge in their field, and moreover, it is a powerful recruiting tool as well. Cockburn and Henderson (1998) further expanded on these ideas and proposed that, at least in pharmaceutical industry, it may be necessary not only to hire the best people and to reward them on the basis of their ranking in the public rank hierarchy, but also to encourage them to be actively connected to the wider scientific community. They found the "connectedness" to be significantly correlated with firms' internal organization, as well as their performance in drug discovery. The estimated impact of "connectedness" on private research productivity implies a substantial return to public investments in basic research. This idea is supported also in another important stream of work which shows that in the case of biotechnology, both rates of firm founding and of new product introduction are related to the connections of the companies to "star" university scientists (Zucker and Brewer, 1994). This research will be summarized and the phenomenon of star scientists analyzed in detail within the following section.

1.2.4 Knowledge spillover mechanisms

1.2.4.1 Knowledge properties, codification and localization tendencies

Knowledge that spills over is considered to be a public good, which means that it is not depleted when shared, once it is made public others cannot easily be excluded from its use and thus it is freely available to those wishing to invest in searching for it (non-excludability), and the incremental cost of an additional user is nearly zero. Knowledge is inherently non-rival in its use, which means that it may be exploited by more than a few users at the same time (Breschi and Lissoni, 2001b). Knowledge developed for any particular application can therefore have economic value in very different applications. The creation and diffusion of knowledge are likely to lead to spillovers and increasing returns (Griliches, 1979).

The distinction between tacit and codified knowledge plays a central role in the literature on knowledge spillovers. Tacit knowledge is "subconsciously understood and applied, difficult to articulate, developed from direct experience and action, and usually shared through highly interactive conversation and shared experience" (Archer and Wang, 2002). It cannot be easily transferred because it has not been stated in an explicit form and its transfer is extremely sensitive to social context. Therefore, the diffusion of tacit knowledge requires the existence of a community of people connected by social links and sharing a common cultural background (Lissoni, 2001). All knowledge for which "a codebook" is available can be classified as codified (Cowan and Foray, 1997). Codified knowledge, on the other hand, can be more precisely and formally articulated. It is described as general and abstract, because understanding it may require high education levels and some personal contacts, even though no common social background is necessary (Lissoni, 2001). Consequently, codified knowledge is easily transferable outside its context of generation, and it can be transmitted through information technologies and infrastructures over long distances, across organisational boundaries and within complex networks at very limited cost and high speed (Cowan and Foray, 1997). The codification of knowledge is a central concept in processes of knowledge dissemination, transfer and retention.

It is also necessary to make a distinction between knowledge and information. Information can be easily codified and has a singular meaning and interpretation. By contrast, knowledge is vague, difficult to codify and often randomly recognized. While the marginal cost of transmitting information across geographic space is invariant thanks to the telecommunications revolution, the marginal cost of transmitting knowledge, and especially tacit knowledge, rises with distance (Audretsch, 1998). Knowledge codification is the process of conversion of knowledge into messages which can be then processed as information. It is actually a transformation of knowledge into information (Ancori *et al.*, 2000). The codification process entails high initial fixed costs, but allows agents to carry out certain operations at very low marginal costs; it is a knowledge transformation into some systematic form that can be communicated at low cost (Cowan and Foray, 1997).

The concept of the tacit knowledge helps to explain the tendency of innovative activities to be concentrated in space. A greater geographic concentration of innovators could be expected if technological knowledge has a tacit nature and cannot be codified through plans, instructions or scientific articles. This type of knowledge can be learned only by everyday practice and use of technology, and informal personal contacts are therefore necessary for its transmission. The use and transfer of new, non-codified knowledge becomes the key to successful development especially when a technology is in the early stages of its life-cycle, because then the knowledge is often very complex and ever-changing. The more the knowledge base of an industry is simple and well codified, the less important is geographical concentration for innovators. Nevertheless, this also probably means that the technology has reached its maturity, and a smaller number of significant innovations could be expected (Baptista and Swann, 1998).

1.2.4.2 Knowledge spillover mechanisms

The literature lists several types of links between firms and the scientific network and consequently several modes for knowledge transfer. The mechanisms facilitating the knowledge spillovers were identified as scientific research published in scientific journals and patent documents, informal contacts and meetings, human capital either embodied in students graduating from university or other workforce mobility, spin-offs from university research and star scientists. Section 3 will analyze specifically the available manifestations of scientific research, which are patents and scientific articles, with a special focus on their use as econometric indicators measuring various aspects of innovativeness. The current section will survey the literature regarding all the other mentioned mechanisms of knowledge spillovers.

- **Informal contacts**

Firms located in clusters usually share common values, which are so important that the firms form their own cultural environment, within which they are linked by specific informal relations in a complex mix of cooperation and competition. In the standard notion of localized knowledge spillovers it is argued that these informal relations, social links and meetings between employees of local firms and university scientists are the main vehicles for knowledge exchange and a common spillover mechanism. Dahl and Pederson (2004) studied the role of informal networks in the development of regional clusters. They confirmed that informal contact between employees in different firms is one of the main means of knowledge transfer between firms in a cluster. Their paper examines empirically the role of informal contacts between engineers in a specific cluster and concludes that the engineers share even quite valuable knowledge by informal contacts. This confirms that informal contacts represent an important channel of knowledge diffusion. These contacts are also suggested to be an efficient way to get relevant valuable feedback while experimenting and testing different technological paths in clusters of horizontally related firms. Maskell *et al.* (2002) suggest that the experimenting firms can easily monitor, discuss and consider the paths taken by the other firms, and in this way learn from the success and failure of others. By comparing different solutions, selecting, imitating, and adding their own ideas they efficiently participate in a continuous learning process. Maskell *et al.* also analyzed the evolution of informal contacts over time. They claim that the creation of informal networks of contacts involves several phases, starting from relations and a transfer of knowledge between two individuals and ending in the formation of entire networks. Development of routines and conventions during repeated interactions leads to the decrease of costs of future interactions, makes the relationship more stable, brings more trust and mutual understanding and facilitates further informal contacts and interactions.

Several authors examined knowledge diffusion through informal channels within the more formal mechanism of information trading. Von Hippel (1987) defined informal information trading as “an extensive exchange of proprietary know-how by informal networks of employees in rival and non-rival firms”. He analyzed informal know-how trading through the framework of a “Prisoner’s Dilemma” and explained both the presence and absence of informal trading of know-how between rivals in terms of maximizing their profits. Von Hippel argued that employees provide information to colleagues from other firms with the expectation of the benefit of receiving valuable information in return, either immediately or in future. Whether an employee reveals the information depends on the competitive value of the information for his company, availability of alternative sources of that information and on the proximity of the information to a domain in which the involved firms compete. According to von Hippel, informal know-how exchange between rival or non-competing firms is the most effective form of cooperative R&D when the value of the know-how is too small to justify an explicit negotiated agreement to sell, license or exchange. Schrader (1991) found that the participation of the employees in informal information trading networks has **a** positive impact on the economic performance of the firm. He recommends that firms therefore should not discourage such transfers, but should instead attempt to make their boundaries more penetrable. In addition, he claims that information trading also promotes innovativeness of the firm. The firm can participate in the trading and acquire valuable information externally only as long as the benefits outweigh the costs for a trading partner. This forces the company to keep up with technical change and to support internal technology development in order to be able to keep interest of the trading partner. Hence, internal technology development and information trading are not substitutes, but rather complements.

The results of empirical investigations of informal information trading of both von Hippel (1987) and Schrader (1991) confirm that firms in the US steel minimill industry routinely trade proprietary process know-how, sometimes even with direct rivals. Their findings also show that the external contacts are important information source for the employees. According to Schrader (1991), 85% of the employees reported that, at least once during the year before the survey, they had been asked by a colleague working in another firm for some specific technical information.

- **Mobility of human capital**

Another important source of knowledge spillovers is human capital. There are two mechanisms facilitating the knowledge spillovers embodied in the workforce. The knowledge is either embodied in the students graduating from university and then transferred from academia to industry, or in the highly qualified workers, which while changing their jobs spill the knowledge over among the firms.

The employment of university graduates have been confirmed as one of the most important channels for disseminating knowledge from academia to industry by Dasgupta and David (1994), Varga (2000), Schartinger *et al.* (2001) and others. Moreover, the importance of this knowledge transfer mechanism for localization of the firms near universities has been proved. Saxenian (1994) argued that spatial proximity to universities can generate positive externalities that can be accessed by the firm through spillover mechanism of human capital, whereas Schartinger *et al.* (2001) claimed that the amount of university educated human capital is one of the major factors influencing firm location. It is explained by Audretsch *et al.* (2005) that proximity offers the possibility of linking students to industry more efficiently, by providing industry and students a pre-employment experience with each other. The authors' findings showed that universities with a high output of students tend to generate more knowledge-based startups.

The mobility of skilled workers is suggested to be a major mechanism through which technical and market knowledge flows locally (Breschi and Malerba, 2001; Breschi and Lissoni, 2001b) and one of the most important mechanisms of knowledge spillovers (Andersson and Ejermo, 2003). Workforce mobility also plays a significant role in the localization of companies, as it had been argued by Keeble (1988) that the biggest determinant of high-tech industry location in Britain is the spatial distribution of highly qualified labour, and its residential preferences. Saxenian (1994) highlighted the benefits of the high annual turnover rate among skilled personnel in Silicon Valley, which in the early 1990s was approximately 20–25%. She argues that by repeatedly changing jobs these scientists, engineers and technical workers substantially contribute to the creation of technology spillovers. In their study focused on the semiconductor industry Almeida and Kogut (1999) examined the role of the mobility of the highly qualified technical workers in the innovative process. The authors study the localization of patents coming from the semiconductor industry and their results confirm that mobility of engineers have an effect on the pattern of citation. Fosfuri and Ronde (2004) built a model of cumulative innovation, where technology spillovers arise endogenously through labour mobility. Their model predicts that in industries where clustering is driven by technology spillovers, labour turnover is high and skilled workers receive, other things being equal, higher wages. These findings were also confirmed by Zucker and Darby (1996b) in their study on the patterns of innovation in the evolution of the biotechnology industry.

- **Company spin-offs**

This section deals with spin-off companies and describes a specific case of the labour mobility which arises when new uncommercialized knowledge serves as a source for generating entrepreneurial opportunities.

The knowledge spillover theory of entrepreneurship was introduced by Audretsch (1995). It states that as investments in new knowledge increase, entrepreneurial opportunities will also increase, because new firms will be started from knowledge that has spilled over from the source producing that knowledge. Specifically, when new economic knowledge cannot be easily transferred to established firms, often because of organizational factors, the holder of such knowledge will start a new firm. The reason behind this spin-off firm creation is the effort of the worker to appropriate the potential economic value of his knowledge through innovative activity (Audretsch, 1998). This is how Audretsch explains that the small or new firms can exploit knowledge created by expenditures on research in universities and on R&D in large corporations and how these companies are able to generate innovative output even if they are undertaking a generally negligible amount of investment in their own R&D. Audretsch and Lehmann (2005) extended this theory when they proved that the knowledge spillover theory of entrepreneurship has a spatial component. Entrepreneurial activity that results from investments in new knowledge will be spatially localized, because the start ups tend to cluster within close geographic proximity to the knowledge source.

A large majority of new high-tech firms were founded as spin-offs from university research or from other firms. For example, the study of Beaumont (1982) showed that more than 90% of the initial studied locations were found within 40 miles of the previous employer of the founder. This is evidence that the previous accumulation of firms in a region provides it with a self-reinforcing advantage in attracting new entrants.

Link and Scott (2005) quantified university spin-off formations into a university research park. Their study analyzed the determinants of the formation of university spin-off companies within the university's research park. The authors found that the formation of the university spin-off companies is more common in older parks, in the parks that are associated with richer university environments, in the ones that are geographically closer to their university and that have a biotechnology focus.

- **Star scientists**

Compared to the previous sections, much more research has been done on the star scientists as an important link between academia and science and a common spillover mechanism. The research in this category is however frequently focused on biotechnology, where the phenomenon is the most apparent. Star scientists in biotechnology are for the purposes of the research defined by Zucker and Darby (1996b) as the scientists with more than 40 genetic sequence discoveries or 20 or more articles reporting genetic sequence discoveries by 1990.

According to Zucker *et al.* (1998a), the majority of the scientists have very low productivity. Most of the scientific output is typically produced by the top 1% or 2% of all scientists working in a specific area. The star scientists are extraordinarily productive, but they account only for 0.8% of all the scientists listed in GenBank through 1990. Nevertheless, they are the authors of 17.3% of the published articles, meaning that their productivity is almost 22 times higher than the average GenBank scientist (Zucker and Darby, 1996b). It is therefore considered logical to focus on the scientific elite, their collaboration with the industry and the localized effects it creates.

The evidence found in the literature shows that the relationship between scientists and firms is symbiotic, as it contributes to the success of both star scientists and science, and the success of the companies and their commercial objectives. Zucker *et al.* (1998a) shed some light on the cooperation between the stars and the companies in the biotechnology industry. Locally linked star scientists provide access to and information about discoveries and advise the firm concerning their bioscience research. The results of Zucker *et al.* (1998a) show that for all three identified measures of firm performance (number of products in development, number of products in market and employment growth) the collaborative research (evidenced by coauthored publications) has a significant positive effect on the firm's performance. Moreover, they claim that the number of star-firm collaborations powerfully predicts success: for an average firm, five articles coauthored by academic stars and the firm's scientists imply about five times more products in development, 3.5 more products on the market, and 860 more employees. However, Zucker and Darby (1996b) reported that the importance of the stars for the company is much lower in the later stages of the development when the new techniques have already diffused widely. Moreover, the cooperation of the company scientists with the star scientists outside their organization is less desirable if the value of the research in question is high. Zucker *et al.* (1996c) also relate the collaboration network structure in biotechnology to the value of the information in the underlying research project: the more valuable the information, the more likely the collaboration is confined to a single organization. As the expected value of research increases, star scientists are more likely to collaborate with scientists from their own organization. Diffusion of discoveries to other scientists decreases as the share of within-organization collaboration increases.

The positive effect of the collaboration between the stars and the companies is also reflected in the higher scientific productivity of the stars. Zucker and Darby (1996b) suggested that stars with commercial ties publish at higher rate before, during and after those ties. Moreover, scientific articles by stars collaborating with or employed by firms have significantly higher rates of citation than articles written by pure academic stars or other articles written by the same stars before or after the collaboration. The authors showed that the presence of just one more affiliated star about doubles the expected citations received by an article. This could be due to the fact that star scientists receive more resources from the biotechnology enterprise and also do the work that is more highly cited while working for or with a biotechnology firm. In addition to that, it was shown that the citations to star scientists increase for those who are more involved in commercialization by patenting. In other words, their research showed that the scientists with patents are generally more widely cited than the scientists without patents, and affiliated scientists are more cited than linked scientists who in turn are more cited than untied scientists. Zucker *et al.* (1998a) confirmed these results, and in addition they argued that those stars affiliated with firms are very different also in their patenting activity compared to unaffiliated university stars. Their results show that 50% of affiliated stars have patented discoveries versus only 15.6% of the university stars. The patenting of discoveries by stars is an indication of expected commercial value of their discoveries.

The importance of the geographic proximity and the geographic linkages among scientists and biotechnology firms are often explored in the literature. The creation of the geographically bounded networks among university-based scientists and the companies is explained by Zucker *et al.* (1998a). Star scientists in biotechnology, who are initially typically employed by universities, appropriate much more benefits from their research than the employing university itself. If their research is potentially significantly successful, they create a spin-off company in order to appropriate the economic value of this research through entrepreneurial activity. After they become involved in commercial applications of their inventions, these star scientists often retain their university affiliations and remain within commuting distance of the university, thus creating the localized effects of university research.

Star scientist was found to be a principal determinant of the location of new biotechnology enterprises. Zucker and Brewer (1994), Zucker and Darby (1996a) and Zucker *et al.* (1998b) provided considerable evidence suggesting that the timing, location and the success of new biotechnology firms is primarily explained by the presence at a particular time and place of scientists who are actively contributing to the basic science.

Audretsch and Stephan (1996) further examined the extent to which the firms and university-based scientists involved with the firms are located in the same region. They conclude that the relationship between the locations of a biotechnology firm and a university scientist is shaped by the potential economic knowledge residing in that scientist and the role that she or he plays in working with the firm. University-based scientists provide three key functions to biotech firms: first, they facilitate knowledge transfer from university laboratories to the firm, which is (given the tacit nature of knowledge in biotechnology) facilitated by face to face contact, and thus it requires geographic proximity. The other two primary functions are signalling the quality of the firm's research to both capital and resource markets, and helping chart the scientific direction of the company. These two functions however do not require geographic proximity and therefore the scientists are less likely to be local. Audretsch and Stephan (1996) in fact found that approximately 70% of the links between biotechnology companies and the university-based scientists are nonlocal in nature. For example, the involvement of older successful scientists with many citations is likely to be also nonlocal, because they are more likely to be known outside their local network than nonpublishers. Moreover, mature scientists with strong reputations have even the drawing power to attract firms to locate near them.

This section has discussed the research regarding the phenomenon of star scientists originating mainly in the United States. There are not many studies concentrated on Canada. It is one objective of this thesis to identify the star scientists in Canadian biotechnology and to examine their role in the collaboration network.

1.2.5 Critique of the localized knowledge spillovers theory

Not only Krugman (1991) doubted that knowledge spillovers are not geographically constrained, but he also argued that they were impossible to measure, because "knowledge flows are invisible; they leave no paper trail by which they may be measured or tracked". Nevertheless, this work has already presented the results of many authors who tracked them and measured their intensity. Breschi and Lissoni (2001a; 2001b) published very critical surveys regarding the research that has proved the existence of localized knowledge spillovers. They complained that authors who claim to prove the effects of localized knowledge spillovers do not test for them specifically, but assume the existence, and if they obtain significant effects from their regressions, then they force their interpretation upon the data. Many of the results however could have been explained by many other effects related to agglomeration or externalities. Breschi and Lissoni believe that authors make logically strange steps when they outline the theory for their research.

They also do not agree with the notion of automatically associating localized knowledge flows to pure knowledge externalities (Bresch and Lissoni, 2001a, 2001b). Breschi and Lissoni suggested that what might appear at first as pure knowledge externalities are actually pecuniary externalities, which are mediated by economic mechanisms (for example the labour market or firm networking) and what might appear as involuntary knowledge externalities are actually well-regulated knowledge flows across firms or between research institutions and firms, that are managed with deliberate appropriation purposes. In fact, even before Breschi and Lissoni's critical paper, there were some voices protesting against the common assumption of knowledge spillovers being pure knowledge externalities. For example, Geroski (1995) argued similarly that what standard methodologies, data sets and concepts proved to be pure externalities will turn out to be, on more careful scrutiny, knowledge flows that are mediated by market mechanisms, which influence local firm's innovation opportunities indirectly, via pecuniary, rather than knowledge, externalities. Zucker *et al.* (1998a) discovered that market mechanisms are the most important facilitator of knowledge transfer in the Californian biotechnology sector. They pointed out that universities, star scientists and firms are usually connected through a contractual system, and thus associated with pecuniary externalities, not spillovers.

It is also emphasized that tacitness may in fact not induce spillovers, but instead contribute to natural excludability. Breschi and Lissoni (2001a; 2001b) observed that much of knowledge transmitted from universities to firms has nothing to do with the public results of basic science, but consists of consultancy services to firms. Rather than providing innovation opportunities, such knowledge transfers may enhance the customer firms' appropriation capabilities. Zucker *et al.* (1998b), Zucker *et al.* (1998a) and Zucker and Darby (1996b) believe that scientific discoveries vary in the degree to which other can be excluded from making use of them. If the techniques for replication involve much tacit knowledge and complexity, and they are not widely known (as is the case of biotechnology), then the degree of natural excludability is high. Natural excludability leads to the embodiment of certain knowledge and techniques in individuals. Under these circumstances, the scientists who make key discoveries (superstars) tend to enter into contractual arrangements with some existing firms or start up their own firm in order to extract the supra-normal returns from the fruits of their intellectual capital.

Several papers (Bresch and Lissoni, 2001a, 2001b; Cowan *et al.*, 2000; Breschi and Malerba, 2001) criticized also the concept of tacitness as an intrinsic inherent property of scientific knowledge. They argue that tacitness rather refers to the way knowledge itself is transmitted and reflects the relative understanding capability of those who communicate, not the specific means of communication. It is suggested that tacit messages can be exchanged at long distances even through very formal means of communications, as long as the level of mutual understanding of those who exchange it is similar. Technical or scientific knowledge is highly specific and its jargon differs from the jargon of the broader social community. The ones who understand it are the members of closed, restricted, but geographically dispersed "epistemic community", within which the tacit messages can be easily transmitted even if knowledge links take place among agents located far away in space. On the other hand, physical proximity does not imply epistemic proximity. The authors claim that tacitness can prevent many local actors from understanding the content of scientific or technical messages. Therefore, knowledge may be inaccessible to most of those who are located nearby its sources. In his case study on Brescia mechanical firms, Lissoni (2001) confirmed that knowledge does not flow freely within the boundaries of the cluster, but circulates within a few smaller "epistemic communities". Each of these communities is centered around the mechanical engineers of individual machine producers, and reaches some selected number of suppliers' and customers' technicians. Physical distances among members of each community vary a lot, but he confirms that even local messages may be highly codified. Moreover, Lissoni also argues that public laboratories and universities are usually not part of these small epistemic communities.

Many theories regarding localization of knowledge creation and diffusion get undermined if the geographical and cognitive proximities are decoupled. In the theory of localized knowledge spillovers, it is argued that the local informal relations, social links and meetings between employees of local firms and university scientists are the main vehicles for knowledge exchange and a common spillover mechanism. However, this assumption is criticized from several points. For example, Hoen (2001) claims that although firms prefer a location near a knowledge institution, they hardly interact with this institution. Lissoni (2001) observed that the informal channels of communication are also not common for sharing knowledge with competitors. Instead, he suggested that inter-personal communication links are much more fruitfully used for sharing knowledge with customers, which is not a spillover mechanism. Schrader (1991) argued that the close social ties and friendship do not play any significant role in raising the likelihood that two engineers will share knowledge and according to von Hippel (1987) any knowledge sharing is likely to involve only exchange of small ideas, while more strategic knowledge is unlikely to be disclosed. Moreover, Lissoni (2001) adds that in epistemic communities the engineers usually remain loyal to the firm and the knowledge exchanged within cluster is thus very general. The firms in clusters are not homogenous and many specialize in very narrow market niches, outside which the firm-specific knowledge is not directly useful. Not much specific knowledge is thus diffused through informal contacts within a cluster. Breschi and Lissoni (2001a) suggest that the members of the community with informal arrangements on sharing knowledge are constrained by the reciprocity obligations, which can act as an exclusionary device. These obligations may make the community members reject the internal contacts and may push them outside their clusters to search for externalities there. Many community members thus end up excluded from the flow of externalities. Prevezer (1997) confirmed that there are many other networks (for example alliances or collaboration) which cross local boundaries and are a method of absorbing information spillovers without having to be situated in the same location.

As a practical example of the doubtful importance of the informal relations within a cluster it was Saxenian (1994), who when comparing two successful regional agglomerations in Silicon Valley and Route 128, pointed out the great differences with regard to the character of the informal contacts within the two clusters. In Silicon

Valley, informal contacts between individuals are important, mutually beneficial, and detailed technical and market information is thus widely exchanged. In the Route 128 case, however, informal contacts are few, because the culture discourages networking and the exchange of knowledge and work-related problems. This shows that the informal cultural environment which is rich of social relations is probably not always important mechanism for knowledge flow and certainly not the prerequisite for successful innovation. Prevezer (1997) even states that evidence that such local social networks are important is anecdotal. Similar conclusion was drawn by Wolfe and Gertler (2004), whose findings do not provide any convincing proof of the direct, non-market interaction and knowledge sharing between local firms in the same industry.

Not only firms seem to be hesitant to share the information, but it is argued that the companies may have incentives to systematically avoid that valuable knowledge spills over. Zucker *et al.* (1996c) observed that whenever discoveries have significant value, whether as pure science or as a commercial product, behaviour has often systematically excluded potential competitors from access to that information. One way to avoid the spillovers is relocation of the whole company to an isolated area, as suggested by Fosfuri and Ronde (2004). They proposed that when product market competition is intense, firms might try to locate in distant areas in order to minimize technology spillovers and preserve their competitive advantage. Consequently, the presence of technology spillovers might turn out to be a reason against industrial clustering. There even has been some work providing recommendations to the companies on how to design their organization to avoid that valuable knowledge spills over to competitors through workers' mobility (Ronde, 2001). The research of Fosfuri and Ronde (2004) focused on the protection of trade secrets and its effect on clustering, spillovers and firms profits. They found that secret protection does not affect clustering. Trade secret protection based on punitive damages is usually beneficial for the company's profits and stimulates clustering. However, trade secret protection that prevents technology spillovers from arising reduces the profits of the firms, because although firms will cluster, technology spillovers do not materialize.

1.2.6 Diversification and specialization externalities

Two types of externalities are usually recognised to play a major role in the process of knowledge creation and diffusion (Glaeser *et al.*, 1992): specialisation externalities (Marshall, 1890; Arrow, 1962; Romer, 1986), which operate mainly within a specific industry, diversity externalities (Jacobs, 1969) which work across sectors and competition externalities (Porter, 1990). Marshall (1890) observes that industries specialise geographically, because proximity favours the intra-industry transmission of knowledge. Jacobs (1969) believes in diversity as the major engine for fruitful innovations, because "the greater the sheer number of and variety of division of labour, the greater the economy's inherent capacity for adding still more kinds of goods and services" (Jacobs, 1969, p.59). A closely related debate concerns competition externalities (Porter, 1990). Porter argues that local competition rather than monopoly favours growth and the transmission of knowledge in specialised geographically concentrated industries.

On the one hand, Marshall (1890), Arrow (1962), and Romer (1986) put forward a concept, which was later formalized by Glaeser *et al.* (1992) and became known as the Marshall-Arrow-Romer (MAR) model. This model claims that the concentration of an industry in a region promotes knowledge spillovers between firms and facilitates innovation in that particular industry within that region. Knowledge externalities between firms, however, only occur among firms of the same or similar industry, and thus can only be supported by regional concentrations of the same or similar industries. It is consequently also assumed that there cannot be any transmission of knowledge spillovers across industries. Glaeser et al. (1992, pp.1127) further argue that "local monopoly is better for growth than local competition, because local monopoly restricts the flow of ideas to others and so allows externalities to be internalized by the innovator." The MAR model therefore perceives monopoly as better than competition as it protects ideas and allows the rents from innovation to be appropriated. These intra-industry spillovers are known as localization (specialization) externalities, Marshall or MAR externalities. In this thesis Marshall or MAR will be used indistinctively.

Jacobs (1969), on the other hand, argues that the most important sources of knowledge spillovers are external to the industry within which the firm operates. Since the diversity of these knowledge sources is greatest in cities, she also claims that cities are the source of innovation. Her theory emphasizes that the variety of industries within a geographic region promotes knowledge externalities and ultimately innovative activity and economic growth. A science base, which facilitates the exchange of existing ideas and generation of new ones across disparate but complementary industries, represents the common basis for interaction. The exchange of complementary knowledge across diverse firms and economic agents thus facilitates search and experimentation in innovation. Jacobs sees diversity rather than specialization as a mechanism leading to economic growth. Therefore, a diversified local production structure gives rise to urbanization (diversification) externalities or Jacobs externalities. A further argument in her thesis concerns competition which is more desirable for growth of cities and firms as it serves as a strong incentive for firms to innovate and hence speeds up technology adoption.

A third type of externality refers to Porter's (1990) argument, also associated with Jacobs[1], that competition is better for growth. Porter also argues that knowledge spillovers occur mainly within a vertically integrated industry, hence agreeing with the Marshallian specialisation hypothesis in identifying intra-industry spillovers as the main source of knowledge externality.

---

[1] Although Jacobs does not formally discuss the effect of competition on growth, the concept is associated with this "school of thought". She is referring to the competition of new ideas rather than in the product market.

MAR and Jacobs agree that there are geographical effects of the agglomeration of firms, but that is as far as it goes. They disagree on the effect of concentration, MAR (and Porter) arguing that knowledge spills over from firms of the same industry, while Jacobs makes the case for variety of industries. The two school of thoughts also disagree on the effect of diversity, Jacobs arguing that knowledge spills over across industries while MAR (and Porter) specifically argue against this. MAR and Jacobs hypotheses also differ in the effect that local competition has on knowledge spillovers and growth Jacobs (and Porter) favour a more competitive environment as conducive to growth while MAR would argue that such an environment is not conducive to innovate as the risks of idea leakages to others are too high.

As a consequence, the question as to which of the Marshall-Arrow-Romer (MAR) or Jacobs externalities is the most beneficial to growth or innovation is rather complex. Whether diversity or specialization of economic activities better promotes technological change has been the subject of a heated debate in the economic literature. It is one of the objectives of this thesis to investigate the reasons behind the inconsistent results of the literature.

1.3 Innovation networks

1.3.1 Collective invention

Allen (1983) examined the British blast furnace industry in the 19th century and proposed that interactions of a collection of firms produce “collective inventions”. The key to understanding a phenomenon of collective invention is in the exchange and free circulation of knowledge and information within groups rather than in the inventive efforts of particular firms or individuals. Owners of blast furnaces shared information (publishing it and presenting it in trade association meetings) regarding the technical and economic properties of their recent furnaces. This produced the discovery of a positive relationship between furnace height and production levels. The open knowledge sharing resulted in fast rates of innovation and rapid productivity growth in blast furnace operation. A large number of other historical examples were afterwards documented in the literature (for examples see Lamoreaux and Sokoloff, 1997). Some examples of the wide informal knowledge trading between engineers in competing minimill firms in the US steel industry (von Hippel, 1987 or Schrader, 1991), of knowledge sharing in a cluster of wireless communication firms in Denmark (Dahl and Pedersen, 2004) or an open knowledge sharing culture in Sillicon Valley (Saxenian, 1994) have already been discussed. There are also interesting modern examples of how collective invention can generate significant amount of knowledge and wealth, as the World Wide Web or the development of public domain (open source) software.

The ideas of collective invention are convenient for describing the dynamics of knowledge diffusion through networks and clusters. Collective invention is characterized by high invention rates and fast knowledge accumulation created by disclosure of information between competing agents. It is driven by exchange and circulation of knowledge and information within networks formed by groups of socially connected individuals (Dahl and Pedersen, 2004). When this is the case, the structure of the network over which transmission of information takes place may be vitally important to the performance of the industry (Cowan and Jonard, 2003).

1.3.2 Network structure analysis

One of the most important features of collective invention is the sharing of information among a broad, typically localized, group of agents. As suggested earlier, the result of this sharing is affected by the network structure over which communication takes place, because it influences the extent of diffusion and thus the innovative potential of the firms. The following section has as an objective to explore this important network aspect.

Social network analysis is the mapping and measuring of relationships and flows between people, groups, organizations, computers or other information/knowledge processing entities. The nodes in the network are the people or groups, whereas the links show relationships or flows between the nodes. Social network analysis provides both a visual and a mathematical analysis of complex human systems (Krebs, 2006).

The methods of social network analysis are commonly used to analyze the way innovators or innovating companies are interconnected. Social network analysis is based on an assumption of the importance of relationships among interacting units (Wasserman and Faust, 1994). Within the research community which investigates the innovation networks it is widely presumed that two innovators, who have worked together on at least one patent or one scientific article, will keep in touch afterwards in order to exchange information or to share some knowledge assets. Similarly, it is presumed that the companies that sign collaboration agreements or jointly own patents are and will remain in important relationships. The patent documents, bibliometric data and the alliance databases are thus frequently exploited to map the complex web of social ties among innovators.

1.3.2.1 Research on knowledge flows in innovation networks

This section provides a summary of empirical studies exploring the structure of two distinct kinds of networks – network of innovators and inter-firm collaboration networks. A special attention is given to the indicators of knowledge flows used in the studies. Afterwards, recent advances in the theoretical simulation modeling of the innovation networks are presented.

- **Networks of innovators**

Network of innovators is an inter-personal network of individual innovators, who collaborate and exchange knowledge for the production of innovations and scientific knowledge. These are the inventors and scientists working at the universities, in research centers or industrial R&D departments. There is usually no formal agreement among the researchers; however, they frequently take part in the development of a patent or the creation of scientific article. Co-inventorship of a patent as evidenced by the patent documents and co-authorship of an article in scientific journal are thus used as the common proxies for knowledge flows between the innovators in order to build their networks.

The network of scientists, whose links are established by their co-authorship of scientific articles, may be the largest social network ever studied (Newman, 2001a). Newman (2001a) was the first (to my knowledge) using four computer databases of scientific papers in physics, biomedical research and computer science to construct networks of collaboration between scientists in each of these disciplines and to study a variety of statistical properties of these networks to describe the network structure. In his subsequent papers (Newman, 2001b; Newman, 2001d), Newman continued his research on the scientific networks, exploring the variety of nonlocal network properties, such as typical geodesic distances (shortest path between scientists through the network), and measures of centrality, such as closeness and betweenness. He also introduced a more efficient algorithm to calculate betweenness and suggested a measure of the strength of collaboration, which can serve as a weight to the collaboration networks. Newman (2001c) then examined empirically the time evolution of scientific collaboration networks in physics and biology.

Breschi and Lissoni (2003 and 2004) and later Balconi *et al.* (2004) presented the idea of how to construct the network of collaborative relationships linking Italian inventors using data on co-inventorship of patents from EPO. They constructed bipartite graph of applicants, patents and inventors. Using this graph, they could derive various measures of social proximity between cited and citing patents. They also calculated geodesic distance, degree of centrality and betweenness centrality to characterize the innovation network. Cantner and Graf (2006) proposed to build the networks of innovators based on technological overlap, which is a measure of closeness of the technological field of two scientists. They also describe the evolution of the innovator network of Jena, Germany using the information on scientific mobility. Singh (2005) inferred collaborative links among individuals using social proximity graph, which he also constructed from patent collaboration data. Many other researchers mentioned later in this section (for example Mariani, 2000, Ejermo and Karlsson, 2006; Bresch and Lissoni, 2003; Gauvin, 1995 and Fleming *et al.*, 2006) adopted the co-inventorship of patents as an appropriate device to derive maps of social relationships between inventors and to build their networks. Based on interviews with inventors, Fleming *et al.* (2006), however, warned that patent co-inventorship links differ significantly in their strength and information transfer capacity. Also, since their decay rates vary greatly, a substantial number of old ties remain viable even if the relation does not exist anymore.

- **Inter-firm collaborative networks**

The network of innovators is distinguished from the inter-firm collaboration network, which consists of a set of companies that are involved in the research collaborative partnerships with other firms. Inter-firm collaborative partnerships represent all the forms of research collaboration among companies and could be evidenced by the existence of strategic alliances, various collaborative research agreements or joint patent ownerships. In order to build the network of strategic alliances, Verspagen and Duysters (2004) used the CATI database of officially registered alliances as a source of information. Also Schilling and Phelps (2005) made use of the information on publicly created strategic alliances to examine the influence of the structure of industry-level alliance networks on firm innovation (measured by patent counts). Other encountered indicators of relationships between the agents in the innovation networks are various collaborative research agreements. Gambardella and Garcia-Fontes (1996) built innovation networks based on EU research contracts, using the official relationships within the contracts to build the links, while Orsenigo et al. (2001) used R&D collaborative agreements to analyze the structural evolution of the collaborative network in pharmaceutical R&D. Moreover, the joint ownership of patents (co-assignees in the patent document) is also considered a sign of the inter-firm cooperation.

- **Theoretical simulation studies**

There are also theoretical simulation studies, in which researchers build models of innovation networks to simulate knowledge diffusion through the network. Cowan and Jonard (2003) have developed a model of knowledge diffusion and studied the relationship between the structure of the network across which knowledge diffuses and the distribution power of the innovation system. Cowan *et al.* (2004) have continued with the simulation study of knowledge flows and compared the mean knowledge growth under different network architectures (ranging from the highly clustered to the one that has no spatial structure). In order to capture the observed practice of informal knowledge trading proposed by von Hippel (1987) and Schrader (1991) and discussed in section 1.2.4, Cowan and Jonard (2004) modeled knowledge diffusion as a barter process in which agents exchange different types of knowledge only if it is mutually profitable. They examined the relationship between network architecture (characterized by different levels of path length and cliquishness) and diffusion performance. Morone and Taylor (2004) identified the limitations of Cowan and Jonard's model (2004) and improved it by introducing a network structure that changes as a consequence of interactions. They investigated the dynamics of knowledge diffusion and network formation. Finally, Cowan *et al.* (2005) modeled the formation of innovation networks as they emerge from bilateral decisions. They developed a model of alliance formation and examined the nature of the networks that emerge under different knowledge and information structures.

1.3.2.2 Properties of innovation networks

- **Properties of the network of innovators**

The results from the abovementioned research studies are summarized in this section, which focuses on the various properties of the innovation networks. Apparent differences in collaboration patterns between the subjects under study were observed. The characteristics of the network structures differ depending on whether they contain purely industrial or also academic researchers. Balconi et al. (2004) observed that networks of inventors within industrial research are usually highly fragmented. On the other hand, the networks constructed by Newman (2001a) were very clustered, but since he based them on scientific co-authorship it could be assumed that these were mainly academic networks. Newman (2001b) also observed that for most scientific authors the majority of the paths between them and other scientists in the network go through just one or two of their collaborators. This could be in agreement with Balconi *et al.* (2004) who found that academic inventors that enter the industrial research network are, on average, more central than non-academic inventors - they exchange information with more people, across more organizations, and therefore play a key role in connecting individuals and network components. Academics also have a tendency to work within larger teams and for larger number of applicants than non-academic inventors (Balconi *et al.*, 2004). Newman's findings (2001a) added that the scientists with larger numbers of collaborators were usually researchers in experimental disciplines, compared to those in theoretical disciplines. Specifically researchers in experimental high-energy physics have a substantially higher average number of collaborators per author than in any other field examined (Newman, 2001a).

Newman (2001c) showed that the probability of a pair of scientists collaborating increases with the number of other collaborators they have in common, and that the probability of a particular scientist acquiring new collaborators increases with the number of his or her past collaborators. Nevertheless, Cantner and Graf (2006) did not find relation between previous and present cooperations with the same partners, suggesting that collaborations are not persistent (in the studied region). Former collaborations are also found to be determinant of the future success. Cowan *et al.* (2005) claimed that previous collaborations increase the probability of a successful collaboration and Fleming *et al.* (2006) argued that an inventor's past collaboration network will strongly influence subsequent productivity.

Cantner and Graf (2006) studied the dynamics of the innovation network and concluded that it is directed towards an increasing focus on core competencies of the local innovation system, meaning that innovators on the periphery exit and new entrants position themselves closer to the core of the network. This implies that a critical mass of innovators is necessary for a specific technology to survive within a local system. Moreover, an increasing specialization of the system should be expected.

Important results of Cowan and Jonard (2003) and Cowan *et al.* (2004) showed that the existence of network structure can significantly increase the long-run knowledge growth rates. The architecture of the network over which innovators interact influences the extent of diffusion and thus the innovative potential of the economy.

To summarize, the networks of inventors which are composed mainly of academic inventors are usually highly clustered and very central. Academics also work in larger teams. The networks of industrial inventors are more fragmented, less central, and are composed of smaller teams.

One of the objectives of this thesis is to study the collaboration networks of Canadian biotechnology and nanotechnology inventors. The aim is to compare the network architecture of both industry sectors and to identify its role in the communication efficiency.

- **Properties of inter-firm collaboration networks**

The typical feature of the inter-firm collaborative networks is that they *are very sparse.* Gauvin (1995) describes his network as a loose meshing of a large number of organizations indirectly connected by few direct links. Cummings (1991) explains that forming and maintaining alliances is expensive and requires time. If the ties are not well maintained they will diminish with time. It is, however, suggested (Angel, 2002) that large firms and firms located in major urban centres are more likely to enter into technology development partnerships, which was confirmed also by Mariani (2000) and Gauvin (1995) who showed that multinational corporations are to a much higher degree engaged in external cooperation. This agrees with the research of Singh (2004), whose results suggest that there are significant bi-directional knowledge flows between multinational companies and their host countries.

The inter-firm collaboration networks are also decentralized. Most of them have several greater components rather than one giant component observed in networks of innovators (Newman, 2001a), or one single dominant firm that connects all other firms in the network (Gulati and Gargiulo, 1999).

Alliance networks tend to be highly clustered, which means that some groups of firms will have more links connecting them to each other than to the other firms in the network (Schilling and Phelps, 2005).

There seem to be great inter-country differences in the collaboration patterns in inter-firm networks. The results of Mariani (2000) suggest that even though the European inventors in chemical industry usually do collaborate (75% of the patents have at least 2 inventors), they prefer to keep the collaboration within the same institution (Only about 8% of all examined patents had multiple assignees). According to Gauvin (1995), the percentage of joint ownerships within Canadian patents is even lower (4% in 1989). The UK research seems to be very open to external collaboration, but German chemical companies prefer in-house research (Mariani, 2000). The US, the country with the largest share of all the patents, is also the one with fewest joint ownerships of intellectual property. Japanese firms to a larger extent engage in cooperation, and their coalitions are also more international and more often cross-sectoral. Otherwise, cooperation generally involves the partners of the same country and the same industry (Gauvin, 1995).

In sum, even though there are great difference between the collaboration patterns among distinct countries and cultures, in general, inter-firm networks are sparse, decentralized and highly clustered.

1.3.3 Brokers and gatekeepers

Over the last two decades there has been an emerging interest in the role of intermediaries in the innovation process. Brokers are either individuals or organizations who "facilitate transactions between other actors lacking access to or trust in one another" (Marsden, 1982, p. 202). By enabling the flow of resources between otherwise unconnected groups the brokers assume an important role in innovation networks and thus received plenty of attention from the research community. Howells (2006) reviews the long history of research on the contribution of brokers to the development and commercialization of technology carried out within a number of different research fields.

One of the most widely acknowledged works is Burt's (1992) theory of structural holes, which describes how the firms embedded in sparse networks of disconnected partners gain efficiency and control benefits. A "structural hole" is a gap in the flow of information between subgroups in a larger network. A firm occupying many structural holes has an advantage over competitors, because it has an easier access to information (due to many non-redundant contacts) and a greater control over the flow of information between disconnected partners. The empirical research has confirmed the power that the brokers gain due to their network positions. In his analysis of the US manufacturing sector, Burt (1980) finds that profit margins of the firms in industries that are situated between disconnected sellers and buyers are higher. Fernandez and Gould (1994) show that organizations which occupy brokerage positions in the national health policy domain are more likely to have greater perceived influence. Similarly, Burt (2004) finds that individuals who span structural holes in an organization gain substantial social capital (compensation, positive performance evaluations, promotions, etc.). Stuart *et al.* (2007) show that biotechnology firms acting as brokers have higher chances to make profitable alliances with downstream partners. Burt (2007) nevertheless points out that the brokerage benefits are dramatically concentrated in the immediate network around a broker, but the benefits are much reduced in case of second-hand brokerage (transfer of information between people with whom a broker has only an indirect connection).

Winch and Courtney (2007) describe the role of innovation brokers, which are organizations specifically founded to undertake intermediary role – to transfer the knowledge between the sources of new ideas and the users of those ideas in innovation networks. But Hargadon and Sutton (1997) stress that brokering is more than just transferring knowledge. A broker also serves as a repository of knowledge, which allows him to recombine existing ideas from various resources and to generate solutions to the problems in other industries. However, brokered ideas seem to be less likely used in future creative efforts. Fleming *et al.* (2007) illustrate how collaborative brokerage can aid in the generation of an idea but then hamper its diffusion and use by others.

Obviously, the brokerage role is quite varied, and brokers can facilitate transactions in a number of distinct ways. In their classification of brokerage roles, Gould and Fernandez (1989) identify types of brokers based on the network configurations that result when a broker connects two otherwise unassociated partners. Among other possible roles, a broker can also act as a gatekeeper. It was Allen (1967) who first identified certain industrial researchers in an organization as key persons in the innovation process, because they gather, process and transfer information from internal and external processes. These individuals were labelled gatekeepers. Allen's study has initiated the creation of various gatekeeper concepts.

The role of gatekeeper has been studied at two levels of analysis: cluster and firm. At a cluster level, the gatekeepers are characterized as leading firms that search for non-local knowledge, transmit it into the region and thus link the region with the outside world (Morrison, 2008). Leading firms can act as gatekeepers not only due to the well-established external contacts, but also due to their superior knowledge base, technological resources and capabilities that make them better equipped to absorb new knowledge and facilitate its diffusion throughout the cluster (Malipiero *et al.*, 2005). The absorptive capacity of the gatekeepers is also at the heart of the research of Lazaric *et al.* (2008) who propose the way of its effective realization, while Boschma *et al.* (2007) study the impact of the local network positions of the firms and their connectivity to the non-local firms on their innovative performance. However, it is not only private firms that assume gatekeeper functions, but also research universities and cooperative R&D institutions (Steiner and Ploder, 2007). Public research organizations have been even suggested to serve the functions of a gatekeeper to a higher degree than private actors (Graf, 2008).

Research on gatekeepers has less often been carried out at the firm level of analysis, where the exchange of information between the individuals within one company is usually the main focus of the study. It has been shown (Allen, 1977; Tushman and Katz, 1980; Katz and Tushman, 1981) that the total performance of the R&D system in the firm is in fact critically dependent on a few key individuals – the gatekeepers, because they provide a linking mechanism between the company and its external environment. Harada (2003) investigates the knowledge transforming function of the gatekeepers and suggests the presence of other key individuals within R&D organizations – knowledge transformers.

Overall, there is a lack of research on the individual gatekeepers carried out in a more global context. The studies mentioned above focus either on the role of the gatekeepers-firms within a cluster or on the role of gatekeepers-individuals within a company, but the whole national network of these individuals together with all their intra-cluster and inter-cluster connections has not been taken into consideration. Therefore this thesis poses two main research questions: Who are the key individuals which enable the nurturing of clusters with fresh external knowledge? How can these gatekeepers be identified in the national network of inventors and how can their importance for the cluster and for the country be evaluated? It is one of the objectives of this thesis to provide answers to these questions.

CHAPTER 2

RESEARCH QUESTIONS, OBJECTIVES AND METHODOLOGY

The research in this thesis is organized around the four already introduced themes: knowledge externalities in regions and clusters, Canadian high technology clusters, collaboration and networks, and Canadian biotechnology prominent inventors. This chapter first introduces the research questions, then explains concrete research objectives and describes the general organization of the thesis clarifying which chapter will satisfy each objective. Finally, it explains some methodological issues.

2.1 Research questions

The clusters lie at the heart of the thesis. The first issue studied concerns the role of the industrial composition of a cluster: The clusters could be highly specialized, which is the case when most of the companies and supporting services in the cluster belong to one leading industry, but they can be also highly diversified, meaning that many diverse industries are represented in one cluster. The question how the composition of economic activities in the cluster influences the growth of the region has been asked by many researchers before who ended up with quite inconsistent answers, and so it comes again to be tackled in this thesis. This work however shifts away from the usual research questions - Do firms in the specialized or diversified clusters perform better? Does specialization or diversity more promote innovation? – and asks why it is that the academic community still has not reached a consensus regarding this issue.

The main focus of the thesis is however on Canadian biotechnology clusters. The key research questions posed here are: Does Canadian biotechnology innovation take mainly place in clusters? Where exactly are these clusters? Who are the inventors in

Canadian biotechnology clusters? What about the companies, universities and governmental research labs? What is their role in the Canadian biotechnology innovative process? Do these companies and/or institutions collaborate together when creating innovations? Who are their main cooperation partners and where these reside? Does the collaboration mainly take place inside cluster, between clusters or is it international? Does geographical distance play a role in the selection of the collaboration partner?

Another array of research questions concerns individual biotechnology researchers: Are all of the researchers in the cluster of the same importance in communicating the gathered knowledge further? Are there some prominent researchers instrumental in supplying the clusters with external knowledge? How could these gatekeepers be best identified and their importance evaluated? Are there some inventors which produce significantly more innovation than others? How to identify these star scientists? What are their positions in the collaboration network? Are these highly productive inventors the same individuals as the ones who nurture the clusters with the knowledge originating outside the region or are the roles of star scientists and gatekeepers separated?

All of these questions relate to the Canadian biotechnology clusters. However, the thesis also provides a comparison with another field that is also of high importance for Canada – nanotechnology. Both of these sectors are high tech, are characterized by a great amount of tacit knowledge and have a high propensity to patent. The idea behind comparing the findings of two fields is to investigate how the various circumstances, particular conditions and features of each field are reflected in the collaboration patterns and the network structures. Here, the key research questions asked are: Is the evolution of biotechnology and nanotechnology innovation in Canada similar? And what about the collaboration patterns in both fields? Do inventors have similar or distinct preferences in the selection of their collaboration partners? Are the knowledge networks through which the innovation in clusters is created comparable? What are the differences and similarities?

The following section formulates the main objectives which create guiding structure of the thesis and which enable answering the questions posed above.

2.2 Research objectives

Objective 1: Explore the role of knowledge externalities in regions/clusters

- Investigate the reasons behind the inconsistent results of the literature regarding the impact of specialization and urbanization externalities on the economic performance of the firms in regions/clusters

Objective 2: Describe the creation of innovation in Canadian biotechnology clusters

- Identify, analyze and characterize Canadian biotechnology clusters
- Describe the impact of the intellectual property rules at Canadian universities on their propensity to patent in biotechnology

Objective 3: Investigate the role of geography in the collaboration

- Determine the collaboration pattern in Canadian biotechnology innovation
- Determine the impact of geographical proximity on the selection of the collaboration partners

Objective 4: Compare the biotechnology and nanotechnology innovation in Canada

- Compare the collaboration characteristics in biotechnology and nanotechnology innovation
- Compare the innovation network architecture and its role in the communication efficiency within the biotechnology and nanotechnology clusters

Objective 5: Study the gatekeepers in the collaboration network

- Establish the way to identify the gatekeepers who enable the nurturing of biotechnology clusters with fresh information originating outside and how to determine their relative importance as procurers of external knowledge for the cluster or for Canada

Objective 6: Identify the star scientists and examine their role in the innovation network

- Establish a method of identification of the Canadian biotechnology star scientists
- Identify the network positions of the star scientists

2.3 General organization of the thesis

The way these objectives were attained is described in the course of 6 chapters. Chapter 3 deals with the issue of knowledge externalities and explores its role in regions or clusters. It describes two types of externalities, which play a major role in the process of knowledge creation and diffusion: specialization (Marshall-Arrow-Romer or MAR) externalities which operate mainly within a specific industry, and diversity (Jacobs) externalities which work across sectors. Therefore, the economic performance of cluster is either supposed to be promoted by the concentration of a particular industry in a cluster (MAR) or it is the diversity of industries in a region which should be most beneficial to growth and innovation (Jacobs). Whether specialisation or diversity of economic activities better promotes development in the region has been the subject of a heated debate in the economic literature. During the literature review a great inconsistency in the results of research papers arguing and providing evidence for the support of or opposition to either theory has been encountered. This chapter attempts to find the reasons behind the inconsistent results as it provides a census of the papers that have dealt with the MAR-Jacobs dichotomy and searches for the similarities between the various studies. Moreover, the threshold at which either theory becomes dominant from the point of view of the level of industrial aggregation, of spatial agglomeration, etc, will be identified. By exploring the various roles of knowledge externalities in regions and clusters Chapter 3 thereby satisfies the ***Objective 1***.

The next chapter deals with the clusters at a much more concrete level. It identifies, analyzes and describes Canadian biotechnology clusters with a special focus on the innovation creation. A profile description for these clusters in terms of patenting quality and quantity, the nature of biotechnology activities, the properties of the assignees and their propensity to collaborate will be made. A crucial role of the publicly funded research in Canadian biotechnology will be highlighted and an importance of the well developed intellectual property policies and functioning technology transfer offices at Canadian universities identified. Chapter 4 thus fully attains the goals stated under the ***Objective 2***.

Chapter 5 continues to tackle the issue of the Canadian biotechnology clusters; however the light is shed here on the collaborative activity within and among the clusters. In keeping with ***Objective 3*** the collaboration pattern in Canadian biotechnology innovation will be described and then the role of geography in the selection of collaborative partners investigated.

The main goal of Chapter 6 is to compare the biotechnology and nanotechnology innovation in compliance with the ***Objective 4***. The comparison of the evolution of biotechnology and nanotechnology patenting and the main collaboration characteristics in biotechnology and nanotechnology clusters will be made. The principal objective of this chapter is however the investigation of the local collaboration in the cluster-based subnetworks. The structural properties of the biotechnology and nanotechnology subnetworks will be examined and compared and their efficiency in knowledge diffusion and innovation creation discussed.

The following two chapters remain within the realm of Canadian biotechnology innovation, but the focus is shifted from the clusters to the individual inventors. The two chapters concentrate on the key individuals in the innovation process. Chapter 7 studies the gatekeepers – the inventors who by bridging over the geographical space (cluster) and technological space (network) occupy most favourable places in the innovation network to fulfil the role of the suppliers of fresh information originating outside their own cluster. In accordance with ***Objective 5*** the method which enables the identification of the gatekeepers in the collaboration network is proposed. The method is based on the determination of the relative importance of each inventor as procurer of external knowledge either for the cluster or for Canada.

Star scientists are at the heart of Chapter 8. First, a new method for their identification is proposed. These prominent inventors in Canadian biotechnology are found by taking into consideration either only patent quantity (stars scientists), or both the patent quantity and quality (*QQ-star inventors*). Moreover, the network approach is adopted again and the positions of the star and *QQ-star inventors* are studied in a complex net of innovative collaborations. Finally, this chapter also involves the examination of the overlap between the gatekeepers and the stars or *QQ-stars*. Chapter 8 thus meets ***Objective 6.***

Chapter 9 then summarizes all the conclusions and makes recommendations for future research.

2.4 Methodology

2.4.1 Patents and their use in innovation research

"A patent is a temporary monopoly awarded to inventors for the commercial use of a newly invented device" (Trajtenberg *et al.*, 1997). There are many advantages for the use of patents to evaluate innovative activity. The information on patents can be easily obtained; together with R&D they are the most easily accessible kind of data (Andersson and Ejermo, 2003). Patents are by definition related to inventiveness and are based on an objective and only slowly changing standard (Griliches, 1990). They are richer, finer and have a wider coverage than for example R&D expenditure (Trajtenberg, 1990) and thus can greatly increase the precision in the statistical analysis (Andersson and Ejermo, 2003). They are also continuous in time. Patents are in fact the only manifestation of inventive activity covering virtually every field of innovation in most developed countries and over long periods of time (Trajtenberg, 1990).

However, there are important limitations of patents as indicators of innovation. First, not all inventions are patentable. For a patent to be granted, the innovation must be novel, non-trivial, and has to have commercial application. Second, not all inventions are patented. Patenting is a strategic decision and for inventors of the patentable inventions it may be preferable not to apply for patents and rely on secrecy instead. According to the survey of Statistics Canada (2002), 66% of the firms which protected their intellectual property have chosen to do so by confidentiality agreements, whereas patenting was adopted only by 39% of respondents. Also, Jaffe and Trajtenberg (1996) recognized that much of research that is performed at both universities and government laboratories never results in patents. Thus, patentability requirements and incentives to refrain from patenting limit the scope of measures built on patent data (Griliches, 1984, Trajtenberg *et al.*, 1997).

Moreover, there are differences in patent office practices across time, technological area and countries (Hall *et al.*, 2001). The differences in procedures and resources of various patenting offices imply also the difference in the average quality of a granted patent across countries and periods. Mansfield (1984) argued that the value and cost of individual patents vary enormously within and across industries as well. Griliches (1990) confirmed that the propensity to patent also varies across the industries, and states that the industries with the largest numbers of patents are drugs, plastics, other rubber products, computers, instruments, communication equipment and industrial chemicals. Scherer (1983) mentioned that the propensity to patent is not invariant across a wide range of firm sizes as well. Consequently, Griliches (1990) and Trajtenberg (1990) argued that patents vary tremendously in their technical and economical significance, which makes it dangerous to draw definitive conclusions based on number of patents. Trajtenberg *et al.* (1997) suggested that the number of patents cannot account for the enormous heterogeneity of research projects and outcomes that characterize the R&D process. They identified two sources of this heterogeneity as basicness (originality, closeness to science, breadth, etc.) and appropriability (ability of inventors to reap the benefits from their own innovations) and proposed the indicators that can capture them.

However, patents are still quite popular indicator of innovativeness and have been commonly used in many of the studies discussed previously, for example Jaffe (1989), Kelly and Hageman (1999), Acs *et al.* (2002) or Fisher and Varga (2003). Simple patent count, which is the number of patents assigned over a certain period of time and is measured at the level of firms, industries, countries, etc., is most frequently used as an indicator of innovative output, innovative input or market value of the firms.

Patent count is a commonly used measure to indicate the innovative output (especially in the knowledge production function), however Griliches (1990), Scherer (1991) and Mansfield (1984) have all observed that patent counts measure only an intermediate output in the entire process of producing an innovation. Griliches (1990) and Mansfield (1984) warned that measuring the number of patented inventions is not equivalent to a direct measure of innovative output.

Trajtenberg (1990) also claimed that single patent counts cannot be very informative of the value of the innovative process (innovative output), however he pointed out that simple patent counts in fact reveal more about the input side (reflected in R&D outlays), because it incorporates the differences in efforts in the innovative process (input). Among the major findings of the survey paper of Griliches (1990) was the discovery of a strong relationship between patent numbers and R&D expenditures (relationship between R&D and patents is close to proportional) in the cross-sectional dimension, implying that the patents are a good index of inventive activity across different firms. Patent counts are suggested to be used as an indicator of inventive input (Griliches, 1990). A strong relationship between R&D and the number of patents was also observed by other researchers, (for example Griliches, 1984). However, Jaffe (1986) found that the payoff in terms of patents to a firm's own R&D is higher in technological areas, where there is much R&D undertaken by other firms. This means that the companies in clusters with high R&D intensity will receive more patents per R&D dollar than firms in other areas, implying that the relation between the patent counts and the R&D expenditure is not that strong, and the patent count is a noisy indicator of an innovative input.

Griliches (1981) found a significant independent effect of patents on the market values of firms, above and beyond their R&D expenditures. Nevertheless, there are other studies (Porter, 1998) trying to relate the patent counts to value indicators (for example market value of innovating firms), but they did not show any significant relation. It is argued (Griliches, 1984) that this is caused mainly because patents vary enormously in their technological and economic significance. Pakes (1985) in his research on the relationship between patents, R&D and the stock market rate of return finds no evidence that independent changes in the number of patents applied for (independent of current and earlier R&D expenditures) produce significant effects on market's valuation of the firm. The patent counts thus cannot be considered a reliable indicator of market value.

It could be concluded that even if the patented inventions differ greatly in their quality, inventive output and economic impact, the patent count is a somewhat noisy but valid measure of innovativeness (especially if the patents pertaining to one industry and registered at one patent office are used). In this thesis, patents are used as an evidence that a certain innovative activity has taken place (depending on the kind of analysis it is considered as an activity of an inventor, of an assignee or it is counted per cluster), but also as a proof of a collaboration relationship among all the co-inventors listed in that patent.

2.4.2 Biotechnology database

The patent database used for the empirical analysis regarding the biotechnology clusters is the United States Patents and Trademarks Office (USPTO) database. This is the only patent database which provides the geographical location of the residence for each inventor (unlike the Canadian Intellectual Property Office database (CIPO) or the European Patent Office (EPO)). The use of the USPTO database instead of the CIPO may introduce a bias in the data, but it is expected to be minimal, since Canadian inventors usually patent both in Canada and in the US. For example, in 1998 and 1999 out of all the patent applications submitted by Canadian biotechnology firms worldwide, the majority (36%) was delivered to the USPTO, followed by 28% to the CIPO, 21% to the EPO and the balance of 16% to other offices[2] (Statistics Canada, 2001). The population of Canada is relatively small and as a consequence, building a viable industry based on domestic sales alone may prove difficult. In addition, because of the long development cycles for biotechnology products (typically 10 years for a single product), access to large markets is needed to ensure an adequate return on investment (Strachan, 1995). As a result, Canadian biotechnology firms prefer to protect their intellectual property in the USA. The much larger and easily accessible US biotechnology market offers great potential to Canadian biotechnology firms. An analysis of the Canadian patents registered at the USPTO should hence provide a realistic picture of Canadian biotechnology innovation. Many researchers who study Canadian biotechnology clusters (for example Niosi, 2005 or Aharonson *et al.,* 2004) also use the USPTO database instead of CIPO.

Biotechnology encompasses several different research technologies and several fields of application. A Statistics Canada study (Rose, 2000) has shown that different interpretations of the meaning of biotechnology can result in differences in the results of biotechnology surveys. One of the initial tasks was therefore to select a clear and practical definition of biotechnology. It has been opted to base the USPTO search strategy on the OECD definition of biotechnology, which is based on the group of

---

[2] Note that firms may have submitted patents regarding the same invention to a number of patent offices at the same time. For instance, some patents are registered at the USPTO, CIPO and EPO.

carefully selected International Patent Codes (IPC)[3]. The OECD has carried out an extensive consultation (including the work conducted by Statistics Canada, which shares similar definitions of biotechnology (Munn-Venn and Mitchell, 2005), to develop the definitions of biotechnology techniques, and the validation showed that the definition appears to capture a significant proportion of biotechnology patents. It might not be complete and may include some patents with non-biotechnology techniques, however, errors are likely to be small (OECD, 2005).

An automated extraction program[4] was used to collect the required information[5] from the biotechnology patents. All the biotechnology patents registered before March 31, 2007 were included. According to the above definition, there are around 100 000 biotechnology patents registered at the USPTO. A patent database, which contains all the patents in which at least one inventor resides in Canada and which comprises 3550 patents, has been created. The total numbers of Canadian patents registered at the USPTO each year found by the aforementioned search strategy largely correspond with what other authors have found: for example the results of Statistics Canada (2001) or the study of Rasmussen (2004). Substantial differences in the findings of other researchers have nevertheless been noticed. These were usually caused by the choice of different search strategies such as keywords in the patents' names and abstracts (as in Niosi and Bas, 2001) or by the decision to use a rather narrow biotechnology definition (as in Beaucage and Beaudry, 2006).

---

[3] The OECD definition of biotechnology patents covers the following IPC classes: A01H1/00, A01H4/00, A61K38/00, A61K39/00, A61K48/00, C02F3/34, C07G(11/00, 13/00, 15/00), C07K(4/00, 14/00, 16/00, 17/00, 19/00), C12M, C12N, C12P, C12Q, C12S, G01N27/327, G01N33/(53*, 54*, 55*, 57*, 68, 74, 76, 78, 88, 92).

[4] Thanks to the InovarisQ team.and specifically to Ahmad Barirani for creating the program. It constituted a part of his Master's thesis.

[5] Extracted information necessary for the research leading to this paper includes the patent number and the inventors' names and their addresses.

2.4.3 Nanotechnology database

The nanotechnology data used in this thesis is based on the Nanobank database[6], which is a public digital library comprising data on nanotechnology articles, the USPTO patents and the US federal grants. In this thesis, only the Nanobank patent database has been used.

When working with Nanobank an outstanding number of patents in nanotechnology[7] has immediately been noticed. Nanobank roughly contains 240 000 nanotechnology patents registered at the USPTO between 1976 and 2005, whereas other sources of reference suggest the total number to be much smaller. Surveying the literature, much smaller samples of nanotechnology patents as shown in Table 2-1 have been discovered. The comparison of the results from these studies is not clear-cut. Some of these works do not encompass the complete period of Nanobank (1976-2005), especially the last two or three years which are undoubtedly the most fruitful ones in terms of the nanotechnology patent production. In fact, it is only after 1998 that the USPTO patent applications started to accelerate considerably. Moreover, the range of these numbers is substantial and reflects the complexity of identifying the relevant nanotechnology bibliometric data in general. Nevertheless, these estimates still provide at least a rough idea of the scale of the results from other researchers. None of their counts and estimates is anywhere near the 240 000 USPTO patents present in Nanobank and identified as related to nanotechnology. In the belief that the Nanobank authors probably found a better method for the nanotechnology patent identification the Nanobank content has been scanned. The presence of both nanotechnology relevant and "not so related" patents has however been discovered. The next step hence involved

---

[6] Nanobank ©2007 by Lynne G. Zucker and Michael R. Darby, Los Angeles, CA: UCLA Center for International Science, Technology, and Cultural Policy and Nanobank. The permission to use the release 1.0 (beta-test) was obtained. See the Nanobank database website: http://www.nanobank.org/.

[7] According to the USPTO, nanotechnology patents are those patents, whose subject matter has at least one physical dimension of approximately 1-100 nanometres, and which involve a special property, function or effect that is uniquely attributable to the nanoscale physical size.

consulting the experts in the field in order to test the nanotechnology relevance of the patents in the Nanobank database.

Table 2-1: Comparison of the number of nanotechnology patents found by other authors

Reference	*Patents*	*Reference*	*Patents*
Meyer (2001)	**2 624**	National Science and Technology Council (Bailey, 2003)	**7 000**
Rothaermel and Thursby (2006)	**3 236**	ETC (2005)	**7 004**
Sampat (2004)	**3 748**	Wong *et al.* (2007)	**7 034**
Darby and Zucker (2004)	**3 900**	Huang *et al.* (2007)	**7 406**
Bonaccorsi and Thoma (2006)	**4 500**	Kanama (2006)	**17 200**
Lee *et al.* (2006)	**4 965**	Huang et al. (2007)	**17 544**
Lux Research (2006)	**4 996**	Marinova and McAleer (2003)	**32 000**
Li *et al.* (2007)	**5 363**	Derwent Web of Nanotechnology (2003)	**35 000**
Berger (2006)	**5 000**	Porter *et al.* (2006)	**54 000**
National Cancer Institute (2006)	**6 000**	Huang *et al.* (2007)	**97 509**
Bhaskarabhatla (2006)	**6 000**	Nanobank	**240 000**

A Canadian Nanobank database, which only contains the patents with at least one inventor or co-inventor residing in Canada and which comprises 5076 such patents, has been created. In total 2070 patent abstracts randomly selected in Canadian Nanobank have been sent out for evaluation of their nanotechnology relevance to various nanotechnology experts[8] including mainly professors at universities, but also industrial nanotechnology consultants (see the example of the questionnaire in Appendix A). Evaluations of 391 patent abstracts, out of which 347 were marked by a definitive response (44 patents were marked by an "I don't know" answer), have been received. The distribution of these definitive responses is shown in Table 2-2.

According to the experts consulted, only 4% of patents were evaluated with certainty as being nanotechnology patents and another 24% could probably deal with the nanotechnology related topic as well. Around 34% of patents are probably not nanotechnology relevant and 38% of the patents deal certainly with a completely non-nanotechnology related theme. It could thus be assumed that around 72% of Nanobank

---

[8] Most of them were identified as nanotechnology researchers on the website of NanoQuebec (www.nanoquebec.ca), a non-for-profit organization which plays a major role in planning and shaping nanotechnology innovation in Quebec.

patents, in which at least one inventor or co-inventor resides in Canada, are not (or probably not) related to nanotechnology, while only 28% of them are (or probably are) nanotechnology relevant. Understandably, this is not a rigorous survey but it serves the purpose of giving an indication of what is perceived as nanotechnology and what is not. Further investigation into what should be considered nanotechnology is required.

Table 2-2: Nanotechnology experts' answers to the question "Is this a nanotechnology related patent?"

Positive answer		*Negative answer*	
YES	Probably YES	Probably NO	NO
4%	24%	34%	38%
Total positive 28%		**Total negative 72%**	

Relative to other technology areas, searching for nanotechnology-related patents is complicated, because nanotechnology covers a broad class of disciplines, materials and systems. Academics from one discipline could then potentially misclassify a patent as non nanotechnology if they are not familiar with the specific domain of the patent. Until recently, there has been no formal classification scheme for US nanotechnology patents. In 2004 the USPTO created new classification code for nanotechnology and started classifying the nanotechnology patents retroactively. The patents that use key terms related to nanotechnology were selected and then manually reviewed (NCI, 2006). As of November 2007, the US Class 977 contains 4815 nanotechnology patents; this process is however not finished yet. The US classification system is thus an insufficient tool for identification of the nanotechnology related patents and the keyword search strategies are used instead. However, due to the multi-disciplinary nature of nanotechnology it is very challenging to find and judiciously use appropriate keywords while searching in a patent database. Some of the strategies employed by researchers have been reviewed.

The nanotechnology relevant publications or patents may be found using solely the prefix "nano*", which indeed should identify a great majority of works. Some of the researchers employed this strategy for constructing their databases of nanoscience publications or patents.[9] The most common methodology for the identification of nanotechnology-related patents however consists in using "nano*" as a basic filter in conjunction with other selected keywords and their variations, creating thus a set of unique keywords pertaining to nanotechnology.[10] Some authors[11] performed extensive testing for a substantial number of potential search terms to assess their specificity with regards to nanotechnology. The most elaborate methodologies consist in the formulation of a set of keywords generated using various iterative techniques with relevance feedback.[12]

According to Kepplinger (2004), who is a Deputy Commissioner for Patent Operations and a USPTO patent examiner, even though there are many patents that include terms related to nanotechnology in the patent disclosure, there is currently only a limited number of patents that actually claim a nanotechnology invention based on a text search and manual review. Therefore, most of the researchers apply exclusion terms and various restrictive strategies to filter out the patents which may use some of the keywords without really pertaining to nanotechnology. For instance, phenomena like "self-assembly" or "self-organization", which are keywords present in most search strategies, are not necessarily nano-specific, and methods like "transmission electron microscopy", which is also a frequent search term, could be applied to different fields as well. Huang (2003, 2004 and 2007), who identified the second highest number of the USPTO nanotechnology patents mentioned in the literature, has failed to use exclusion and restrictive terms, which may explain the exaggerated figures (see Table 2-1) obtained particularly for the full-text searches.

---

[9] E.g., Braun et al. (1997), ISI (2002), Tolles (2003), Darby and Zucker (2004).

[10] Majority of the reviewed papers used this methodology: e.g., Bachmann (1998), Meyer (2001), Noyons et al. (2003), Marinova and McAleer (2003), Warris (2004), Heinze (2004), CREA (2005), Sampat (2005), Bhaskarabhatla (2006), Berger (2006), Rothaermel and Thursby (2006), Wong et al (2007), Li et al. (2007) and Huang et al. (2003, 2004 and 2007).

[11] E.g., Porter et al. (2006).

[12] Kostoff (2006), Mogoutov and Kahane (2007), Zucker et al. (2007).

The aim of this thesis was to find a strategy which would allow obtaining the largest possible extent of relevant only data. As in most of the reviewed studies, the prefix "nano" has been used to find the core of the nanotechnology patents in conjunction with complementary keywords that better define the field and extend its borders, but exclusion terms and restrictive conditions have been applied in order to keep non-relevant patents outside. The final search strategy shown in Table 2-3 is largely based on that of Porter *et al.* (2006), but is modified to suit the purposes of this research. The search algorithm on the full text of the patents has been applied to keep it as inclusive as possible and 2493 nanotechnology patents with at least one inventor or co-inventor residing in Canada have been found. When compared these patents with the Nanobank content, only 1442 of them were included simultaneously in both databases. Such a small magnitude of overlapping between the databases is very surprising. It means that 72% of patents in the Canadian Nanobank database contain neither the search string "nano*" nor any other of the commonly used keywords anywhere in the text[13]. One possible explanation was offered by Bawa (2004), a registered patent agent at the USPTO, who remarks that nanotechnology patents often do not use any specific nano-related terminology in order "not to be found" to keep potential competitors at a knowledge disadvantage. On the other hand, some inventors and assignees might incorporate nano-relevant terms only for the sake of marketing their invention or concept even if the inventions are in fact not related to nanotechnology.

Specifically for this research it is extremely important to include only strictly nanotechnology related patents. In contrast, authors whose main interest is in generating the innovation trends or comparing the patent production proportions may find that an overstated total number does not necessarily lead to the wrong conclusions. However, since one of the objectives of this thesis is the construction of social networks, an inclusion of additional actors can significantly alter the network properties. Therefore it

---

[13] Out of all original Canadian Nanobank patents (5076 patents), only 1442 (28%) are also among the results of the current search.

Table 2-3: The final search strategy based on Porter *et al.* (2006)

Search terms	*Search queries*
Nano$ with exclusion terms	nanoa$ OR nanob$ OR nanoc$ OR nanod$ OR nanoe$ OR (nanof$ ANDNOT nanofarad$) OR (nanog$ ANDNOT (nanogram$ OR nano-gram$)) OR nanoh$ OR nanoi$ OR nanoj$ OR nanok$ OR (nanol$ ANDNOT nanoliter$ OR nano-liter$) OR (nanom$ ANDNOT nanomol$) OR nanon$ OR nanoo$ OR nanop$ OR nanoq$ OR nanor$ OR (nanos$ ANDNOT (nanosec$ OR nano-sec$)) OR nanot$ OR nanou$ OR nanov$ OR nanow$ OR nanox$ OR nanoy$ OR nanoz$ OR (nano ANDNOT (nano-gram$ OR nano-liter$ OR nano-sec$ OR nano-meter$ OR nano-metre$))
Quantum terms	“quantum dot” OR “quantum dots” OR “quantum array” OR “quantum arrays” OR “quantum device” OR “quantum wire” OR “quantum wires” OR “quantum computing” OR “quantum well” OR “quantum wells” OR “quantum effect” OR “quantum effects”
Molecular terms	“molecular wire” OR “molecular wires” OR “molecular wiring” OR “molecular switch” OR “molecular switches” OR “molecular sensor” OR “molecular sensors” OR “molecular motor” OR “molecular motors” OR “molecular device” OR “molecular devices” OR “molecular ruler” OR “molecular rulers” OR “molecular simulation” OR “atomistic simulation” OR “molecular manipulation” OR “molecular engineering” OR “molecular electronics” OR “molecular modeling” OR “single molecule” OR “single molecules”
Other terms without delimiters	bionano$ OR biomotor$ OR fullerene$ OR “coulomb blockade” OR “coulomb blockades” OR coulomb-staircase$ OR langmuir-blodgett OR “PDMS stamp” OR “PDMS stamps”
Self-assembly terms limited more inclusively to the molecular environment	(“self-assembly” OR “self-assembling” OR “self-assembled” OR “self assembling” OR “self assembled” OR “self assembly” OR “self-organised “ OR “self-organized” OR “self organized” OR “self organised” OR “directed assembly”) AND (monolayer$ or mono-layer$ OR film$ OR quantum$ OR multilayer$ OR multi-layer$ OR array$ OR molecul$ OR polymer$ OR copolymer$ OR co-polymer$ OR mater$ OR biolog$ OR supramolecul$
Microscopy terms limited more inclusively to the molecular environment	(“atomic force microscope” OR “atomic force microscopy” OR “transmission electron microscope” OR “transmission electron microscopy” OR “scanning force microscope” OR “scanning force microscopy” OR “scanning tunneling microscope” OR “scanning tunneling microscopy” OR “scanning probe microscope” OR “scanning probe microscopy” OR “energy dispersive X-ray” OR “X-ray photoelectron” OR “electron energy loss spectroscope” OR “electron energy loss spectroscopy”) AND (monolayer$ or mono-layer$ OR film$ OR quantum$ OR multilayer$ OR multi-layer$ OR array$ OR molecul$ OR polymer$ OR copolymer$ OR co-polymer$ OR mater$ OR biolog$ OR supramolecul$)
Terms limited more inclusively to the molecular environment	(“quasicrystal” OR “quasi-crystal” OR NEMS) AND (monolayer$ or mono-layer$ OR film$ OR quantum$ OR multilayer$ OR multi-layer$ OR array$ OR molecul$ OR polymer$ OR copolymer$ OR co-polymer$ OR mater$ OR biolog$ OR supramolecul$)
Terms limited more restrictively to the molecular environment	(biosensor$ OR solgel$ OR “sol gel” OR “sol gels” OR “dendrimer” OR “dendrimers” OR “dendron” OR “dendrons” OR “molecular sieve” OR “molecular sieves” OR “mesoporous material” OR “mesoporous materials” OR “soft lithography” OR “soft lithographic”) AND (monolayer$ OR mono-layer$ OR quantum$ OR multilayer$ OR multi-layer$)

was decided to work only with the subset of the Nanobank patents which intersects with the results obtained by the current search as well. After some manual exclusions and obvious additions, the database of 1443 Canadian nanotechnology patents has been created. It was concluded that this final database is the best possible representation of the Canadian nanotechnology.

2.4.4 Social network analysis

From these two databases, biotechnology and nanotechnology innovation networks were created by mapping the collaborations of inventors of particular biotechnology or nanotechnology patents. The concept of social network analysis defined above was employed to create connections between inventors and to construct innovation networks. The social network analysis program PAJEK was used to build the networks from the patent data. An analysis of these collaborative networks enabled the description of their structural properties and allowed the understanding of the collaborative behaviour of the inventors inside or outside Canadian biotechnology clusters.

Since the obtained patent data in both databases span over a period of around 30 years (biotechnology: 1976-2007 and nanotechnology: 1976-2005), it is assumed that once inventors collaborate on one patent they continue to be in contact afterwards and are able to exchange knowledge acquired long after the patent has been granted. In Chapter 6 (see Figure 6-1) it will be shown that a great majority of the collaborations in both biotechnology and nanotechnology took place within a relatively short period of time (last 7-8 years), which explains why this assumption can be made. The time of collaboration thus can be disregarded and all links among inventors in the network can be considered as simultaneously active.

CHAPTER 3

THE LOCALISATION VERSUS URBANISATION CONTROVERSY

This chapter deals with the question introduced in the literature review as to which of the Marshall-Arrow-Romer (MAR) or Jacobs externalities is the most beneficial to growth. Whether diversity or specialisation of economic activities better promotes economic growth and innovation has been the subject of a heated debate in the economic literature. The answer seems to depend on the multiple factors – on the way it is measured, where it is measured, on which industries, at which level of aggregation, etc. This chapter aims to provide a census of the papers that have dealt with the MAR-Jacobs dichotomy[14] (i.e. the regression-based studies providing direct answers in the urbanisation versus localisation debate). The aim is not to try to determine which one of the two concepts provides a more favourable environment for innovation and economic development, but to investigate why it is that the literature still remains relatively

---

[14] Less than half of the articles surveyed include competition or Porter externalities in their analyses, as a consequence, the main focus of the chapter will remain MAR and Jacobs externalities. There have been many other studies which dealt with similar kinds of issues. For example, Rigby and Essletzbichler (2002) note the deficiencies of the most common studies which represent agglomeration economies with very vague proxies (e.g. city size) and suggest to precisely measure three kinds of agglomeration economies: input-output linkages, labour pooling and technological spillovers (following Marshall (1920)). They claim that these agglomeration economies are more precise in their meaning than localization and urbanization economies. Duranton and Puga (2004), on the other hand, do not regard the Marshall's classification of agglomeration economies as a particularly useful basis for taxonomy of theoretical mechanisms, since these agglomeration economies are in fact three sources capturing the same mechanism. They suggest distinguishing theories by the mechanism driving them and propose yet another formulation based on the notions of sharing, matching and learning, which brings the analysis down to a more basic set of variables. Porter (2003) examines the regional performance (wage, employment, patenting), the regional economies and the role of clusters in the US economy. He provides some evidence that specialization of a region in an array of stronger traded clusters boosts regional performance. These are interesting studies, but since the focus of the chapter is on a very narrow concept (i.e. regression-based studies providing direct answers in the urbanization versus specification debate); they will be left out of the further discussion.

inconclusive. This chapter therefore attempts to find the similarities and differences between the various studies in order to draw conclusions on the question.

3.1 Basic results

The phenomenon of knowledge externalities and their impact on economic growth and innovation have attracted a great deal of attention in academic circles. Nevertheless, it seems that the exact spillover mechanism is not yet fully understood and documented. In fact, there is no direct proof of the existence of knowledge spillovers and there probably never will be. A large amount of literature provides empirical evidence in support of the Marshall and Jacobs theories; however, the results of these studies are often mutually conflicting. This section provides a brief survey of these studies and discusses their basic results, while Appendix B contains the summary of the main characteristics, variables, indicators and results from the studies that were examined. The sample of studies is by no means exhaustive.

Table 3-1summarizes the results from the 67 reviewed studies listed in Appendix B, the evidence therein having shown to be in conflict with one another. Around 70% of them claim to have found some proof of existence of Marshall externalities and their positive impact on economic growth or innovative output, while a comparable proportion of the studies (75%) confirm Jacobs' thesis of a favourable influence of diversification of economic activities in a region. Around half of the studies providing support for each theory found uniquely positive results; the other half, however, reported concurrently both positive and negative or non-significant results for various industries, time periods, countries or dependent variables. The situation is similar when looking at the results summary counted by number of variables[15]. Here, as in the remainder of the chapter, each 'variable' used in the models examined to measure externalities (MAR or

---

[15] Many papers use various indicators, or a number of independent variables in their studies. In order to take into account the diversity of results presented within each paper, the number of 'variables' therefore exceeds the number of papers. Please consult Appendix B for more details.

Jacobs) will be counted and linked with each indicator used as a dependent variable. A comparable percentage of variables (57% and 56%) show positive impact. For most of these variables the results are uniquely positive, however, 10% of all variables are found to have both positive and negative or non significant results for different time periods, industries and countries. Moreover, since in most models the positive effects of both kinds of externalities are not mutually exclusive, many researchers have also observed a favourable impact of both Marshall and Jacobs externalities concurrently (30 out of 67 of the studies examined have reported the simultaneous presence of both MAR and Jacobs externalities). Thus, quite a balanced support for both theories is provided by the surveyed studies, hence sufficient evidence exists to claim that both specialized and diversified local industrial structures may promote economic performance of regions. It remains to be found why this may be the case. This chapter represents an attempt at providing an answer.

Although positive evidence for both types of externalities is measured, many of these studies have also detected negative impacts. The score is much less balanced here, because the solely negative influence is observed much more often for the Marshall externalities (in 27% of studies and for 24% of the variables as defined above) than for Jacobs externalities (only in 3% of all the studies and for 7% of all the variables). These findings suggest that regional specialization may hinder economic growth, but diversification is much less likely to induce this negative effect. This may be first related to the lower flexibility of the specialized regions and consequently to their decreased capacity to adjust to exogenous changes, which may prove critical if the main industry in the region declines. In a diversified environment endowed with a wider technological scale, the chances that some new industry will spring out and take the lead is greater. Second, specialized regions may be more vulnerable to lock-in, i.e. closing upon themselves, becoming insular and impermeable, and preventing knowledge and fresh innovative ideas from outside to flow in. The specialized regions tend to become more specialized with time, and thus experience increasingly less external relations than the diversified regions.

Table 3-1: Results summary

Results	Number of studies				Number of variables *			
	MAR		Jacobs		MAR		Jacobs	
Only positive	23	34%	26	39%	51	47%	56	45%
Both positive and negative	**24	36%	**24	36%	***11	10%	***13	10%
Positive sub-total ****	*47*	*70%*	*50*	*75%*	*62*	*57%*	*69*	*56%*
Only non significant	2	3%	15	22%	20	19%	46	37%
Only negative	18	27%	2	3%	26	24%	9	7%
Total	*67*	*100%*	*67*	*100%*	*108*	*100%*	*124*	*100%*

Note. * Each variable used to measure MAR externalities with each indicator used as a dependent variable is counted as a single variable.

** Both positive and negative (or non significant) results found for various dependent variables, time periods, industries or countries within one study.

*** Both positive and negative (or non significant) results found for various time periods, industries or countries for one variable.

**** At least one variable is positive (the sum of only positive and both positive and negative).

Regarding the Porter externalities, only 25 out of 67 studies have attempted to detect their presence. The results of these studies are found in Table 3-2. The Porter theory was most often (14 regressions) supported in conjunction with the Jacobs one, which is also consistent with the Porter's views on competition. Porter however agrees also with the MAR specialization hypothesis regarding the intra-industry spillovers and the two theories were simultaneously supported in 9 regressions. 5 out of all regressions have showed a concurrent support for all the MAR, Jacobs and Porter theses. Since most of the studies did not include the Porter externalities the main focus for the rest of the chapter will remain on MAR and Jacobs externalities only.

Table 3-2: Results for the Porter externalities found in conjunction with Marshall and Jacobs positive results

Porter externalities results	*Number of dependent variables for which positive results were found:*				
	MAR only	***Jacobs only***	***Both***[1]	***None***[2]	***Total***
Positive	4	9	5	2	20
Negative		4	1	3	8
Non significant	4	1	2		7
Total	*8*	*14*	*8*	*5*	*35*

Note:

[1] Number of dependent variables for which both MAR and Jacobs externalities are found.

[2] Number of dependent variables for which neither MAR nor Jacobs externalities are found.

The empirical evidence regarding the nature and magnitude of externalities yields mixed results. This is not surprising, considering that knowledge spillovers are invisible and "leave no paper trail by which they may be measured or tracked" (Krugman, 1991, p.53). The results can be explained by differences in the strength of agglomeration forces across industries, countries or time periods, but also by methodological issues. The remainder of the paper discusses these specific factors and tries to determine the influence of data and the way it is analysed on the likelihood of detecting MAR or Jacobs externalities.

3.2 Indicators for Marshall and Jacobs externalities

The most obvious differences among the studies are the ones associated with the choice of independent and dependent variables. Out of the many independent variables present in the regressions, this thesis will only focus on two categories: local specialization as evidence of MAR economies and local diversity to detect the presence of Jacobs externalities. Some studies, probably constrained by data availability, utilize the same index to measure the impact of both specialization and diversity in the same variable (for example the Hirschmann-Herfindahl index in Loikkanen and Susiluoto, 2002). Authors then may interpret a positive sign (or high values) on the coefficient as evidence of prevailing Marshall externalities and a negative sign (or low values) as a proof of Jacobs economies. This methodology, however, may not be appropriate in some industries because both kinds of economies could be present simultaneously – according to Table 3-6, Table 3-8 and Table 3-12 further ahead in the paper, for 31 (out of 89) dependent variables, evidence of both externalities is found. The two externalities are obviously not mutually exclusive, since specialisation is a particular characteristic of a certain sector within a local system, whereas diversity is a property characterising the whole area. This suggests that testing the two hypotheses separately with different indicators is more appropriate.

3.2.1 Marshall externalities indicators

The location quotient and own industry employment, which together account for 75% of independent variables used in these studies, are the most common Marshall externalities indicators utilised. Other measures encountered are the number of industry plants (either total numbers or relative to plant sizes), several indices based on technological closeness of sectors, measures indicating the share of own industry in a region (measured either by output, R&D investment or industry value added) and other indicators listed in Table 3-3. These indicators are divided into four categories according to whether they measure the:

- Share: indicators based on the relative sizes of the industry, where the proportion a particular industry within the same or other industries in the country, region, and so on, are calculated;
- Size: indicators considering absolute sizes of the industry expressed by employment, number of plants, and so on;
- Diversity: indicators based on industrial diversity using technological closeness of industries, specialization of the science base, and so on;
- Other indicators: not allocated within any of the categories above.

This categorization will be used throughout the article in order to designate which kind of indicator shows a greater number of positive results for Marshall externalities.

Table 3-3: Number of indicators (independent variables) of Marshall externalities

MAR externalities indicators	*Category*	*Only +ve*	*Both* + & -*	*Only −ve*	*Non significant*	*Total*	*Number of studies*
Location quotient (simple or as a proportion of national share)	Share	19	5	16	5	45	35
Own industry employment (total, over area, in innovative or non-innovative firms)	Size	21	5	6	4	36	17
Number of industry plants (total, of minimum sizes)	Size	2			3	5	2
Indices based on technological closeness	Diversity	1			2	3	2
Share of own industry in a region (by output, R&D investment, value added)	Share	1		1	1	3	3
Science base specialization	Diversity	2				2	2
Herfindahl index of concentration	Diversity	2				2	2
Employment in related industries or in provider sectors	Size				2	2	1
Matrix of sectors	Share	1			1	2	2
Autoregressive coefficients	Size	1			2	3	2
Other – share of a firm's innovative activity in an industry, share of own industry in total industry, region's share in national own industry employment, other industry employment, index of production specialization, weighted indices	Share, Size, Diversity or Other	1	1	3		5	5
Total		*51*	*11*		*20*	*108*	

Note. * Both positive and negative (or non significant) results found for various time periods, industries or countries for one variable.

Glaeser *et al.* (1992) first expressed the idea that the degree of specialization may better represent the potential for Marshall externalities than current size of an industry, because it better captures intensity and density of interaction among firms. The location quotient has become widely used for this purpose; it is the most frequently used indicator in the studies reviewed (it is used in 35 studies). The location quotient belongs to the category of indicators based on industrial share, since it represents the fraction of industry employment in a region relative to the national share. The results produced by the regressions that utilised this measure are mostly uniquely positive (19 variables) for the Marshall indicator, however a large number of cases showing negative impacts of specialization (in 16 cases) is found as well. In a number of studies, however, a simpler location quotient is used to measure MAR externalities as a share of a region's employment in an industry. Van Soest *et al.* (2002) has compared the results of the two indices for the same data and concluded that, at least in case of the Netherlands, the relative location quotient (relative to the national share) better captures the impact of Marshall externalities than its simpler - version industry proportion in the region. The relative indicator of specialization controls for the size of industries at the national level, whereas the simpler indicator does not. The evidence from the studies examined shows that simpler measures of MAR externalities are more likely to yield positive results. For instance, the more complex location quotient comparing to the national share provides the vast majority of the negative specialisation effects. Its simpler version very rarely yields negative results.

The location quotient has been criticized as an indicator of local specialization. Ejermo (2005) observed that this measure is very sensitive to the size of the region. Combes (2000b) shows the flaws in the calculation of the location quotient and his corrections of this measure significantly reverse the sign of the relative concentration effect on local growth.

A much simpler measure of the level of local specialization is own industry employment, the most frequently encountered indicator of the category based on the absolute size of the industry. Own industry employment is used in 17 studies (36 independent variables)), mostly with uniquely positive results for Marshall externalities (21 variables), while negative impacts are detected in only 6 cases. Own industry employment is sometimes suggested to be a better proxy for localization economies than the location quotient, because the localization economies arise from the absolute and not the relative size of the industry (for instance Marshall's size of the skilled labour pool).

A region might represent a strong cluster in a certain industry, even if this industry accounts for a negligible share of the region's overall range of activities. It has also been suggested to distinguish between employment in innovative and non-innovative firms in a given industry, because not all employees generate spillovers (Beaudry and Breschi, 2000, 2003) and spillovers are more likely to emanate from firms that also innovate.

Henderson (2003) decomposes own industry employment in a region into the number of plants and the average employment in those plants to discover that it is the number of plants in a region that produces the strongest results. He suggests that localization externalities derive from the existence of companies *per se*, where these companies could be interpreted as separate sources of information spillovers.

The two most commonly used measures of specialization, location quotient and own industry employment, did not produce balanced effects in the regressions. When the results presented in Table 3-3 are compared, it becomes evident that the use of the location quotient yields a greater number of negative results for Marshall externalities than the use of own industry employment. These negative effects emanate almost exclusively from the use of the relative location quotient (relative to the national share), a measure employed almost twice as often as its simpler version. These two location quotient measures contribute equally to the positive results. These positive specialization impacts are also similar for both the location quotient and own industry employment. As a consequence, if one wants to find positive MAR externalities, the simple location quotient, and to a lesser extent own-industry employment, are the way to go. This would tend to favour the argument towards to size of the skilled-labour pool rather than its relative size. Others factors may also influence these results and following the next section on Jacobs externalities, this is what the chapter aims to investigate.

3.2.2 Jacobs externalities indicators

Measures of Jacobs externalities encountered in the reviewed studies are of an even greater variety. The most common is the Hirschman-Herfindahl index (in 38 studies), other industry employment (in 10 studies), Gini index, total local population, total local employment and others. The full list of the diversity indicators and the associated results is presented in Table 3-4 which is a summary of the results presented in Appendix B. Indicators of Jacobs externalities are also divided into categories according to the different focus of the measures. The category based on diversity covers different measures of industrial diversity and specialization, while the one based on market size represents the scale of these urbanization economies and includes various employment or population measures. The category of other indicators is used for those which are not allocated to either of the mentioned groups.

The Hirschman-Herfindahl index is a diversity-based measure and in all of its forms it is the most commonly used indicator (49 variables in 38 studies). The results are split approximately half and half between positive and neutral effects and almost no negative results are obtained. The basic form of Hirschman-Herfindahl index is expressed as the sum of the squares of the shares of employment in a given region and sector with respect to all other industry employment. Other variations frequently encountered are the innovation diversity index based on patent data or industry diversity index based on industry value added data. This Hirschman-Herfindahl index is also presented in modified forms, as inversed Hirschman-Herfindahl index or 1 minus the Hirschman-Herfindahl index. The main drawback of this index is that diversity is measured symmetrically, implying that it does not consider how different or complementary the industrial sectors are, but assumes them to be equally close to one another.

The second diversity-based measure of Jacobs externalities is the reciprocal Gini index, encountered in 13 cases (7 studies), 10 of which yielded uniquely positive signs on the coefficients suggesting the presence of Jacobs economies. The Gini index of diversity is generally used with employment or with patent data.

Other industry employment, which belongs to the category of indicators based on size, is the second most popular indicator of Jacobs externalities, used in 19 cases (10 studies). In 8 cases this indicator does not show any influence of large industrial composition, in 6 cases it shows a negative impact and in only 5 cases it provides some evidence of Jacobs externalities. This indicator does not measure diversity per se but the size of the urbanisation externality. Diversity is implied by the larger size of the employment base in all other industrial sectors. In many of the studies that employ this technique, a measure of diversity is also put in place, so as to account for both the scale and diversity of the urbanisation economies. As a proxy for measuring regional diversity, total employment in the region (also total manufacturing employment or total employment in services) or total population in the region are used as well. In these models, it is assumed that regions with higher population or employment are the ones with more diversified economic structures. These indicators, however, capture rather global urbanization externalities, which are related to local market size, but not to the diversity implied by Jacobs externalities per se, because they derive from the specific industrial composition of the region.

The choice of diversity measure seems to be critical for the result regarding the presence of Jacobs externalities. While the use of other industry employment usually shows negative or no effects from a large diversified region, the Gini index of diversity provides positive findings. The selection of Hirschman-Herfindahl index yields an equal number of positive or neutral results on Jacobs economies.

Table 3-4: Number of indicators (independent variables) of Jacobs externalities

Jacobs externalities indicators	*Category*	*Only +ve*	*Both* + & -*	*Only -ve*	*Non signify-cant*	*Total*	*Number of studies*
Hirschman-Herfindahl index (employment, patent, industry value added)	Diversity	21	6	1	21	49	38
Other industry employment (total, in innovative or non-innovative firms)	Size	2	3	6	8	19	10
Gini index of diversity (employment, patent, science base)	Diversity	10			3	13	7
Total urban area population	Size	6			2	8	6
Total local employment (employment, in manufacturing and in services)	Size	3	2		2	7	6
Share of other industry employment (5 largest, 11 largest)	Size	4			2	6	3
number of active industries in a region	Size				2	2	1
Ellison-Glaeser index	Other				2	2	1
Share of innovations or industries with the same science base	Diversity	2				2	1
Indices based on technological closeness (patents, sectors)	Diversity	1			1	2	2
Related variety	Diversity	1		1		2	1
Share of other industry output	Other		1			1	1
Weighted indices of several elements	Other	1		1		2	1
Other – indices of specialization, diversification, urbanization, Theil index, urban to rural continuum codes, matrix of sectors and other based on expected population, weighted own industry employment and region's employment share	Share, Size, Diversity or Other	5	1		3	9	10
Total		*56*	*13*	*9*	*46*	*124*	

Note. * Both positive and negative (or non significant) results found for various time periods, industries or countries for one variable.

3.3 Performance measures

All the research studies examined can be classified into three categories according to the performance measure under study, which specifies whether the main point of interest is the influence of a specialized or diversified region on economic growth, productivity or innovation. Table 3-5 presents the summary of the dependent variables (performance measures) used to assess these impacts and the number of positive results obtained for each category of independent variables (Marshall and Jacobs externalities indicators). This allows us to investigate where exactly the positive results for each dependent variable come from.

Table 3-5: Performance indicators by category of independent variable and industry classification level

	MARSHALL EXTERNALITIES									*JACOBS EXTERNALITIES*								
Dependant variable	*Indicators with positive results:*				*Total %*[2]		*Industry class.(+ve)*[3]			*Indicators with positive results:*				*Total %*[2]		*Industry class.(+ve)*[3]		
	Share	*Size*	*Diver*[1]	*Total*	*+ve*	*-ve*	*Broad*	*Medium*	*Detail*	*Diver*	*Size*	*Other*	*Total*	*+ve*	*-ve*	*Broad*	*Medium*	*Detail*
ECONOMIC GROWTH																		
Employment (growth or size)	9	9		18	46%	36%	44%	25%	71%	18	10	1	29	74%	10%	81%	100%	50%
Number of new firms (total or per area, proportion)	2	4		6	75%	0%	100%		50%	5	2	1	8	80%	0%	67%		100%
Wage growth (adjusted or not)	2			2	50%	50%	0%	100%			1		1	25%	0%	50%	0%	
Other economic growth: plant size, number of plants (total or per area), of employees per area				0	0%	0%	0%			2			2	50%	0%	50%		
Economic growth sub-total	*13*	*12*	*0*	*26*	*49%*	*26%*	*44%*	*50%*	*64%*	*25*	*13*	*2*	*40*	*70%*	*7%*	*74%*	*67%*	*64%*
PRODUCTIVITY																		
Plant output, output per labour hour, TFP	6	3	1	10	71%	7%	70%	75%			3	1	4	22%	6%	18%	29%	
Valued added growth (or VA over labour)		2		2	67%	33%	50%	100%		2	1		3	75%	0%	100%	50%	
Other: efficiency scores, capacity to export	1		1	2	100%	0%	100%				1		1	33%	0%	0%		
Productivity sub-total	*7*	*5*	*2*	*14*	*74%*	*11%*	*69%*	*80%*		*2*	*5*	*1*	*8*	*32%*	*4%*	*29%*	*33%*	
INNOVATION																		
Number of patents (total or per capita)	3	8	4	15	65%	22%	53%	100%	67%	10	2	1	13	36%	17%	13%	100%	83%
Number of inventions reported by journals	1	1		2	50%	50%	100%		0%	2		1	2	40%	0%	0%		100%
Likelihood of an innovation adoption	1	1		2	50%	0%		50%		1			2	100%	0%		100%	
R&D intensity		1		0	0%	100%	0%			2			2	100%	0%	100%		
Other innovations: number of innovations, of innovators, innovativeness, impact	2			3	75%	0%	67%			1	1		2	50%	0%	33%		
Innovation sub-total	*7*	*11*	*4*	*22*	*59%*	*24%*	*57%*	*75%*	*40%*	*16*	*3*	*2*	*21*	*43%*	*12%*	*16%*	*100%*	*88%*
Total	*27*	*29*	*6*	*62*	*57%*	*23%*	*53%*	*68%*	*56%*	*43*	*21*	*5*	*69*	*53%*	*8%*	*45%*	*64%*	*74%*

Note.

[1] Diversity indicators were grouped with Other indicators

[2] Percentage of the total of indicators, which showed positive or negative results, the remaining percentage showed non-significant results.

[3] Percentage of the positive results per each industry classification level

Table 3-6: Number of positive results (dependent variables) per performance indicator

Dependent variable	*Number of dependent variables with positive results:*				*Total*
	MAR only	*Jacobs only*	*Both MAR and Jacobs*	*None of MAR and Jacobs*	
ECONOMIC GROWTH					
Employment (growth or size)	4	13	12	3	32
Number of new firms (total or per area, proportion)	1	2	5		8
Wage growth (adjusted or not)	2	1		1	4
Other economic growth: plant size, number of plants (total or per area), of employees per area		2			2
Economic growth sub-total	*7*	*18*	*17*	*4*	*46*
PRODUCTIVITY					
Plant output, output per labour hour, TFP	9	2	1	2	14
Valued added growth (or VA over labour)		1	2		3
Other: efficiency scores, capacity to export	1		1		2
Productivity sub-total	*10*	*3*	*4*	*2*	*19*
INNOVATION					
Number of patents (total or per capita)	4		7	1	12
Number of inventions reported by journals	2	2			4
Likelihood of an innovation adoption			2		2
R&D intensity		2			2
Other innovations: number of innovations, of innovators, innovativeness, impact	2	1	1		4
Innovation sub-total	*8*	*5*	*10*	*1*	*24*
Total	*25*	*26*	*31*	*7*	*89*

3.3.1 Economic growth

Most of the research focuses on measuring economic growth (46 regressions in 41 studies), taking the indicator of employment growth as a proxy. Other dependent variables used for this purpose are the number of new firms (total, per area, a proportion), wage growth (adjusted or non-adjusted), plant size, number of employees per firm, number of plants (total or per area) or number of employees per area.

It was expected to find positive results mainly for Jacobs externalities, since the economic growth depends strongly on the level of local demand. Diversified regions with a strong local demand and many intermediaries in the supply chain were supposed to perform better economically. As Table 3-5 shows, Jacobs theory is indeed more often supported (70% of variables with positive results and only 7% with negative) than that of Marshall (49% of positive but 26% of negative results) by these studies. Positive results for Marshall variables came both from share-based and size-based indicators, whereas Jacobs externalities are detected using mainly diversity-based and less frequently by size-based independent variables. There are also 17 regressions showing positive results for both Marshall and Jacobs indicators simultaneously (see Table 3-6).

Employment growth is the most common dependent variable (32 regressions in 30 studies). An overwhelming number of the studies found evidence of some externalities when using this performance indicator, most frequently only Jacobs externalities, while only a few observed uniquely Marshall effects. Favourable results for both these types of externalities are detected simultaneously in many regressions as well. The popularity of this indicator probably stems from the fact that data on total employment are often readily available. It is used when the unit of observation is the firm or the region. In studies at the firm level as opposed to cluster or regional level, the lifetime growth model assumes exponential growth since its creation (Swann *et al.* 1998 for instance). The use of employment growth as an indicator of economic growth is, however, often disputed. The measure of employment growth is based on the assumption that labour is a homogeneous input and that it can freely move across the country. Almeida (2006) suggests, however, that labour is in fact a very heterogeneous input and migration costs differ across countries and periods of time. Cingano and Schivardi (2004) show that a number of other forces are likely to affect local employment determination: a higher unemployment risk against sectoral shocks resulting from sectoral concentration or negative congestion externalities related to the scale of local productive activity may influence mobility and employment choices as well. Moreover, capital and labour have a high degree of substitutability (Paci and Usai, 2005) and the fact that technological change is labour-saving may cause the indicator of employment growth to not properly reflect economic growth. Cingano and Schivardi (2004) show that, within the same sample, if one uses employment growth instead of total factor productivity growth as the dependent variable, the signs for the MAR coefficient are reversed. They claim that these results question the conclusions of most of the existing literature on dynamic externalities. Dekle (2002) reaches similar conclusions about the inappropriateness of employment-based regressions. As an improvement, Combes *et al.* (2004) suggest decomposing local industrial employment into the product of average plant size and the number of plants.

Authors using entry of new firms to the region (measured by the growth of the total number of firms, or the number of firms per area or as a proportion of incumbent firms) as a proxy for regional economic growth (8 regressions in 7 studies) find positive effects of both Jacobs and Marshall economies. Wage growth is used in 4 regressions (in 3 studies) to assess the impact of dynamic externalities on economic growth. MAR variable is found to have a positive impact in half of them, while no positive effects of Jacobs externalities are detected. In order to evaluate the result differences caused by the use of different indicators Glaeser *et al.* (1992) and Almeida (2006) compare the impact of various indicators on wage growth as an alternative measure to employment growth. Glaeser *et al.* (1992) find similar results for both indicators, whereas the signs in the Almeida's regressions are reversed. Almeida (2006) also proposes to use regional adjusted wage growth to account for the heterogeneous character of labour. These obviously call for further investigations.

3.3.2 Productivity

Given the limitations raised by some authors regarding the use of economic growth indicators, researchers have tried to study the impact of the local economic structure on industrial productivity more directly. Productivity-based measures are theoretically closer to the notion of dynamic externalities and may represent some improvement over employment-based measures; the common problem, however, is data availability, since output data (either at firm level or aggregated at regional level) are usually more difficult to obtain. In 18 reviewed studies with 19 regressions (see Table 3-6), the most common productivity indicators used are plant output, output per labour hour, total production factors, value added growth (total or over labour), efficiency scores or capacity to export.

Here the size of the labour pool was considered to be the most important and hence the positive effects of MAR externalities were expected to be detected more often. In specialized regions with bigger labour pool people learn easily from each other, and the absorption of different experiences from people specialized in similar fields contributes to the faster build–up of their skills and thus to their higher productivity. Productivity was not supposed to be influenced much by the size of local demand. As expected, the regressions (see Table 3-5) have more often shown Marshall externalities to be promoting regional economic productivity (variables were positive in 74% and negative in 11% of cases), while the Jacobs theory is supported less frequently (positive only in 32% and negative in 4% of cases). The results for Jacobs were most commonly non-significant. Only 4 regressions have shown positive results for both MAR and Jacobs externalities concurrently (see Table 3-6).

The positive coefficients of the specialization indicators are found using the independent variables of both share-based and size-based categories, while the positive coefficients of the diversity variables originated mainly in size-based and less commonly in diversity-based indicators. Marshall and Jacobs theories are supported concurrently in 4 regressions that use productivity-based dependent variables.

3.3.3 Innovation

The third group of studies (with 24 regressions in 19 studies) attempts to assess the influence of the specialization and diversification of regions on their innovative activity and that of the firms within. The number of patents (total or per capita) is the most frequently selected proxy for innovative output. Other indicators encountered are the number of inventions reported by trade journals, R&D intensity, the likelihood of adopting a particular innovation, the number of innovators, innovativeness or economic impact of an innovation after 2 years.

Innovation could be supported by a great variety of factors within specialized and diversified regions. The effect of both local demand and a labour pool should play role. Indeed, regressions with a dependent variable assessing innovative activity have yielded balanced results for both theories (see Table 3-5), i.e. around half of them in support of Marshall's theory (22 positive variables, which is 59%), while the other half promotes Jacobs' thesis (21 positive variables, which is 43%). Positive results for Marshall variables came both from share-based and size-based indicators, while Jacobs externalities are detected using mostly diversity-based independent variables. There are also 10 regressions with positive results for both Marshall and Jacobs indicators (see Table 3-6).

Patents have long been used as an indicator of innovation, because they are closely related to innovativeness and are based on a slowly changing standard; patent information is also quite easily accessible and of wide coverage. There are nevertheless important limitations to the use of patents as indicators of innovation as summarized by Jaffe and Trajtenberg (2000): "Not all inventions are patentable, not all inventions are patented, and even if they are, they differ greatly in their quality, inventive output and economic impact, making simple patent count quality a noisy measure of innovativeness". To increase the quality homogeneity, Baten *et al.* (2005) use only patents that are being renewed for at least 10 years. Paci and Usai (1999) weight the number of patents with a dimensional variable (by counting patents per capita) to correct for the high heterogeneity in the dimension of the territorial units.

Only 3 studies (with 4 regressions) have utilized the literature-based innovation output method introduced by Acs and Audretsch (1987) to retrieve invention counts. This innovation indicator is considered to be a more direct measure of innovative activity than are patent counts. Innovations that are not patented but are introduced to the market are included in the database, but inventions which are patented but never developed into innovations because they did not prove economically viable are excluded. This innovation count indicator suffers some drawbacks as well: The significance and quality of innovations still vary considerably, the trade journals report mainly product innovations (Feldman and Audretsch, 1999) and the probability to announce a new product in a journal is not equal for all firms and products (Van der Panne and Van Beers, 2006).

* * *

It seems that the performance measure selected as dependent variable has an important influence on the final results. On the one hand, the summarized findings show that Jacobs externalities have a more profound impact on economic growth than Marshall economies. On the other hand, if the influence of the industrial composition on productivity growth is studied, Marshall's theory is more often supported. Only the studies using a dependent variable for assessing the impact on innovative activity have provided balanced support for both theories.

Both absolute size of the industry and its relative size (its share) have a positive impact on the economic performance of a region in the form of Marshall externalities. However, in case of Jacobs externalities, it is mainly the diversity of the industrial base, and to a much lesser extent the size of the regional market (urbanization economies), that promote regional growth. One may ask then what differentiates all these studies. The variations in the results presented here may also emanate from further differences in study construction as will be presented in the following section.

3.4 Other circumstances and conditions of analysis

In this section the effects of specialized and diversified regions will be further studied. The main goal is to analyze positive results of independent variables within each category, for each kind of externality and to investigate the variation in effects which these externalities present under different conditions and circumstances. The specific characteristics which will be considered are the level of analysis (regional versus firm level), different range and types of industries and sectors, geographical units of different sizes and characters, countries and regions, different lengths of periods of observation and variations in the size of firms.

3.4.1 Regional and firm level analyses

There are two possible levels, at which the dependent variable could be analysed. At the regional level, the effects of the externalities on economic performance of the specialized and diversified regions are usually compared. This means that the dependent variable is measured in the industry-region cross-section. At the firm level, causalities are established between externalities and the firm's growth, productivity or innovative performance, i.e. the dependent variable is a performance indicator of the individual firm. The studies that have adopted the firm level approach have the advantage of being able to treat the economic environment as exogenous, while their obvious drawback is firm selection that may bias the results. Table 3-8 shows the summary of all regressions examined grouped according to the level of analysis. The regional level analysis is much more common (63 indicators in 55 studies) and most often focused on regional economic growth. At the firm level (26 indicators in 22 studies), however, it is much more frequent to study the impact of the regional industrial composition on companies' innovative performance. Table 3-7 suggests that evidence in favour of Jacobs externalities is slightly more common at the regional level (65% of positive results for Jacobs versus 54% of positives for MAR), but in support of Marshall's thesis if measured at the firm level, especially if the impact on economic growth or productivity is analysed (in total 62% of positives results for MAR versus 33% of positives for Jacobs).

Positive results for Marshall variables came mainly from share-based indicators at the regional level, and from size-based indicators at the firm level. This suggests that a relative size of an industry (its share in the region) has a positive impact in the form of Marshall externalities on the economic performance of a region, while it is an absolute size of the industry which promotes the growth of the individual firms in a region. Jacobs externalities are detected at both levels using mostly diversity-based independent variables.

Table 3-7: Performance indicators per level of study by category of independent variable and industry classification level

		MARSHALL EXTERNALITIES									*JACOBS EXTERNALITIES*								
Dependant variable		*Indicators with positive results:*				*Total %[2]*		*Industry class.(+ve)[3]*			*Indicators with positive results:*				*Total %[2]*		*Industry class.(+ve)[3]*		
		Share	*Size*	*Diver[1]*	*Total*	*+ve*	*-ve*	*Broad*	*Medium*	*Detail*	*Diver*	*Size*	*Other*	*Total*	*+ve*	*-ve*	*Broad*	*Medium*	*Detail*
	Economic growth	13	9		22	46%	27%	44%	50%	50%	24	12	2	38	75%	2%	74%	67%	88%
Region	Productivity	6	3	1	10	71%	14%	67%	100%		2	2	1	5	29%	6%	23%	50%	
	Innovation	4		3	7	70%	30%	50%	100%	50%	9			9	75%	0%	0%	100%	100%
Sub-total - Region		*23*	*12*	*4*	*39*	*54%*	*25%*	*50%*	*73%*	*50%*	*35*	*14*	*3*	*52*	*65%*	*3%*	*58%*	*77%*	*91%*
	Economic growth		4		4	80%	20%	50%		100%	1	1		2	33%	50%	67%		0%
Firm	Productivity	1	2	1	4	80%	0%	100%	75%			3		3	38%	0%	100%	29%	
	Innovation	3	11	1	15	56%	22%	58%	50%	33%	7	3	2	12	32%	16%	17%	100%	80%
Sub-total - Firm		*4*	*17*	*2*	*23*	*62%*	*19%*	*59%*	*63%*	*67%*	*8*	*7*	*2*	*17*	*33%*	*18%*	*24%*	*44%*	*50%*
Total		*27*	*29*	*6*	*62*	*57%*	*23%*	*53%*	*68%*	*56%*	*43*	*21*	*5*	*69*	*53%*	*8%*	*45%*	*64%*	*74%*

Note.

[1] Diversity indicators were grouped with Other indicators

[2] Percentage of the total of indicators, which showed positive or negative results, the remaining percentage showed non-significant results.

[3] Percentage of the positive results per each industry classification level

Table 3-8: Number of positive results (dependent variables) per performance indicator and the level of study and

Dependent variable		*Number of dependent variables with positive results:*				
		MAR only	*Jacobs only*	*Both MAR and Jacobs*	*None of MAR and Jacobs*	*Total*
	Economic growth	4	17	16	4	41
Region	Productivity	7	2	3	2	14
	Innovation	2	2	3	1	8
Sub-total - Region		13	21	*22*	*7*	*63*
	Economic growth	3	1	1		5
Firm	Productivity	3	1	1		5
	Innovation	6	3	7		16
Sub-total - Firm		12	5	*9*	*0*	*26*
Total		25	26	*31*	*7*	*89*

3.4.2 Industry range and industrial sectors

An important difference in these studies lies in the selected industries. Analyzed data may come from only one industry (as in Beaudry, 2001). The analysis may also consider all the range of industries including non-manufacturing services such as wholesale and retail trade (as in Glaeser *et al.*, 1992, Beaudry and Swann forthcoming), but it is also common to completely exclude services and agriculture from the sample (as in Combes *et al.*, 2004) due to problems of data availability or productivity estimation in services. Furthermore, the methodology may involve an analysis of one manufacturing industry at a time (as in Henderson *et al.*, 1995), which allows to distinguish the roles of either type of externalities in each industry. This approach, however, may not be applicable to all countries, especially in small countries with only a relatively small number of locations where the selected industries can flourish (van Soest *et al.*, 2002). An alternative approach here is to consider only a number of the largest industries of all types in each region (for example the 6 largest industries in each city as in Glaeser *et al.*, 1992), which may *de facto* automatically increase concentration levels in each city. The selected range of industries used for the sample may yield further differences.

In order to determine the factors that may influence the particular suitability of the specialized or diversified region, a more detailed analysis of industrial sectors is carried out. Industries are grouped into four categories: high tech industries, medium tech industries, low tech industries[16] and services. Appendix C presents the list of all the industries compiled according to the reviewed references in which Marshall externalities are found to have positive influence on local performance. Appendix D lists the industries for which Jacobs economies prove to have played a positive role. Not all studies have measured the effects or provided details about separate industrial sectors,

---

[16] The industrial categories were distinguished according to R&D intensity: *high tech* relates to aerospace, consumer electronics, office and computing machinery, semiconductors, pharmaceuticals, biotechnology, nanotechnology, optical and precision instruments; *medium tech* includes electrical and non-electrical machinery, fabricated metal products, motor vehicles, railroad transport equipment, shipbuilding, chemical products, instruments; and *low tech* consists of wood products, furniture, textile and clothing, leather, apparel, food, beverage and tobacco, paper, printing, non-metallic mineral products.

and therefore some references are skipped from the list. Some studies, however, analysed the data for several industries and the positive results are therefore included in each category.

Summary counts of positive results according to industry types are presented in Table 3-9. Surprisingly, the differences among the sectors in regards to the effects which both types of externalities had on regional performance are not striking. It can still be said that externalities probably do play different role in different industries, and that the effects of specialization and diversification economies thus slightly differ across industrial sectors.

Table 3-9: Number of positive results (dependent variables) by industry sector type

	MARSHALL EXTERNALITIES					*JACOBS EXTERNALITIES*			
Industry type	*Indicators showing positive results based on:*								
	Share	*Size*	*Diversity*	*Other*	*Total*	*Diversity*	*Size*	*Other*	*Total*
Low tech industries	11	13			24	10	7	1	18
Medium tech industries	12	14			26	15	10	1	26
High tech industries	7	13		1	21	18	6	2	26
Services	4	1	1	1	7	9		2	11

Low tech, low R&D intensity companies with traditional, more standardized production were assumed to benefit more from the decentralized location in specialized regions, which bears cost advantages and therefore it was expected to detect mainly MAR externalities in low tech sectors. Although not overwhelming, there is some evidence that in low tech sectors Marshall externalities have stronger effects than Jacobs externalities. In 24 studies, the performance of low tech sectors in regions is promoted by Marshall externalities, which is identified by both size-based and share-based independent variables Jacobs' theory is supported in low tech industries only in 18 studies, and this is detected with diversity-based and size-based indicators.

The positive results of medium tech sectors are quite balanced for both theories.

Half of the studies (26 studies) suggested that Marshall externalities are particularly relevant when considering industries grouped into medium tech category. Both size-based and share-based independent variables are used frequently in these studies. The other half (26 studies) proposed that diversified regions are more appropriate for the medium tech industries in conformity with Jacobs' theory. These externalities are identified mainly by diversity-based independent variables.

High tech, R&D intensive companies were expected to prefer to locate in large diversified urban areas, where the cross-fertilisation of knowledge and ideas from outside the core industry, which is so crucial for the high tech breakthroughs, is possible and easily available. The results indeed showed that the high tech sectors slightly favour more diversified regions (in 26 studies), while the effects of Marshall externalities are less pronounced (positive in 21 studies) here. It is mostly size-based indicators which detected Marshall externalities, whereas diversity-based independent variables found most of Jacobs externalities. This would seem to imply that size-based indicators may not be appropriate for innovation measures. Further investigation is therefore needed on this particular point.

The role of externalities varies according to the nature of the sector whether manufacturing or services. Consumer service sectors provide non-tradable goods, which should be produced and consumed in the close proximity of customers. This results in spreading the service activities around and among the customers rather than the concentration of these activities. Business services, on the other hand, greatly benefit from the presence of other sectors located around and are thus concentrated near the firms to which they sell their products. In both cases it was assumed that the location of services is more suitable in the cities (or diversified regions). In accordance with the findings presented in Table 3-9, diversification indeed appears to be the main growth promoter in services (in 11 studies). These positive coefficients came mainly from diversity-based independent variables.

To summarise, it seems that economic performance in both specialized and diversified regions is promoted for all the three groups of industry types. The effects of Marshall externalities are nevertheless slightly stronger in low tech sectors, while the positive impact of Jacobs externalities on regional performance increases with increasing technological intensity. Cross-fertilisations and spillovers may therefore be more useful in high technology sectors. Both relative and absolute sizes of a given industry influence the presence of Marshall externalities for low and medium tech industries, while in high tech sectors, it is mainly the absolute size of the industry which matters. In the case where Jacobs externalities are observed, for all industrial types it is uniquely the diversity of the industrial base, and not the size of the local market, that promotes the regional growth. The size of the industrial base more often then not reflects congestion effects which are detrimental to growth.

The findings of some authors show that the role of externalities of each kind varies in accordance with the maturity of an industry, since old industries might benefit from a different industrial structure than new ones. Henderson *et al.* (1995), for example, find evidence of only Marshall externalities for mature capital goods industries, however, for new high-tech industries, they observe positive effects of both Jacobs and Marshall externalities. These findings are consistent with the industry life cycle model of Duranton and Puga (2001) who show that new industries prosper in diversified metropolitan areas but, when they mature, the production will decentralize to more specialized regions. Also it agrees with the results of Boschma *et al.* (2005), who observe that Jacobs externalities are predominant in the early stages of the industry life cycle, whereas Marshall externalities appear at a later point, and in the end, the specialization will in fact hinder economic growth. The differences in the impacts of the various local industrial compositions during the industry life cycle could be explained by the different needs of the firms during the innovative process. In the initial stage of the innovative process an increased diversity and variety propels the creation of novelty, inventive ideas, creative concepts and radically new designs. When the industry matures and the design reaches a critical mass on the market, the product becomes standardized and the knowledge involved in the innovation process highly specialized. Firms then may greatly benefit from learning from the solutions and mistakes of other firms in the same industry in a region with high concentration of their own industry. Finally, it is the high concentration of the mature industry, which decreases the region's ability to innovate, rejuvenate and restructure, and which inevitably leads the region into a lock-in.

Another factor to consider is the level of industrial classification used for the analysis. An industry could appear as a statistically homogenous entity if 1-digit or 2-digit industrial classification is used, whereas the same industry will look as a diverse assemblage of different activities if the analysis is based on a 6-digit breakdown. Therefore, an analysis of the results per industry class, where broad (1-digit and 2-digit), medium (3-digit) and detailed (4-digit and more) levels of classification were distinguished, has been performed. Not all the studies have indicated which industrial classifications scheme was employed. An educated guess was used for the ones where the classification level seemed apparent but not mentioned, and the ones where the level could not be determined due to the lack of provided information (4 studies with 6 dependent variables studies are thus not included in this analysis) were set aside.

According to Table 3-10, the probability to detect the Jacobs externalities increases with the level of detail of industry classification, whereas it does not have such tendency for the MAR externalities. The highest probability of detection of positive (and the lowest one for negative) results is for the medium level of industrial classification, but is somewhat lower for broad or detailed industrial classification schemes. It has been suggested that completely different indicators may need to be selected to identify the MAR externalities precisely. Some studies which have proposed alternative measures (see Rigby and Essletzbichler, 2002 or Duranton and Puga, 2004) have already been cited. Their arguments are based on the fact that agglomeration externalities do not operate directly on economic growth, productivity or innovation, but are expressions of deeper forces, i.e. output-input linkages, labour pooling effects or localized innovation effects.

However, if staying within the narrow circle of the defined concepts, it can be concluded that in general MAR effects are slightly more prone to show up at the broad level of detail, the probabilities of detecting MAR or Jacobs effects are quite comparable at the medium level and the Jacobs effects will decidedly appear more often when the detailed classification is used. The 3-digit classification could thus be considered as a threshold, at which specialization and diversity are less distinguishable from one another, before which it is specialization and beyond which it is diversity that are favoured.

In order to determine more precisely the diverse effects of the classification level under different conditions the issue was studied even deeper. It was found that when the study is carried out on the regional level, the probability of finding Jacobs externalities is always higher no matter what industrial classification is chosen, whereas on the firm level, it would usually be MAR externalities (see Table 3-7). The selection of a geographical unit however plays a role. The studies which used larger geographical units such as states or provinces and a broadly grained industrial data usually ended up confirming MAR externalities, whereas the studies based on the city level (SMA or MSA in the US) which used detailed industrial data found most commonly the evidence of the Jacobs effects (see Table 3-11). This further confirms the existence of the threshold at the medium classification level. The following section discusses the selection of geographical aggregation further.

Table 3-10: **Number of positive results per industry class by category of independent variable**

Industry classification level	*MARSHALL EXTERNALITIES*							*JACOBS EXTERNALITIES*						*Number of dependent variables for which positive results were found:*				
	Indicators showing positive results:					*Total %*[1]		*Indicators showing positive results:*				*Total %*[1]						
	Share	*Size*	*Diver.*	*Other*	*Total*	*+ve*	*-ve*	*Diver.*	*Size*	*Other*	*Total*	*+ve*	*-ve*	*MAR only*	*Jacobs only*	*Both*[2]	*None*[3]	*Total*
Broad (1-digit and 2-digit)	15	19	1	1	36	53%	24%	21	14	2	37	45%	10%	15	16	17	5	53
Medium (3-digit)	7	4	2		13	68%	11%	11	2	1	14	64%	0%	5	4	6		15
Detailed (4-digit and more)	4	5			9	56%	31%	9	3	2	14	74%	16%	4	5	5	1	15
Total	*26*	*28*	*3*	*1*	*58*	*56%*	*22%*	*41*	*19*	*5*	*65*	*52%*	*9%*	*24*	*25*	*28*	*6*	*83*

Note.

[1] Percentage of the total of indicators (independent variables), which showed positive or negative results, the remaining percentage showed non-significant results.

[2] Number of dependent variables for which both MAR and Jacobs externalities are found.

[3] Number of dependent variables for which neither MAR nor Jacobs externalities are found.

3.4.3 Geographical unit

The selected level of geographical aggregation and the division of the observed territory into regions for the study of geographical specificities could serve as yet another source of possible discrepancy in the results. Table 3-11 and Table 3-12 present the summary of the studies and groups them according to the selected geographical unit. Five different classes of geographical units are observed, with Class 1 being the largest (state or provincial) units and Class 5, the smallest (highly populated areas and cities). Classes 1 and 2 are administrative units, which usually remain unchanged over time and contain the relevant economic market. Class 3 contains all the labour zones, which are the groupings of municipalities, characterised by a high degree of self-contained flows of commuting workers. This makes labour zones economically more homogenous than administrative units. Class 4 represents the smallest postal code level areas, which are usually arbitrary administrative units, not functional economic areas. All these four classes have in common a full coverage of the territory of the country or a selected region, while the areas in Class 5 do not cover the whole surface, but focus only on highly populated areas and cities. Proximity and frequent interactions makes externalities particularly large in a city but, by considering only selected densely populated areas, a large part of the economy is missed.

Table 3-11: Number of positive results per geographical unit by category of independent variable and industry classification level

	MARSHALL EXTERNALITIES									*JACOBS EXTERNALITIES*								
Geographical unit used[4]	*Indicators with positive results:*			*Total %*[2]			*Industry class.(+ve)*[3]			*Indicators with positive results:*			*Total %*[2]			*Industry class.(+ve)*[3]		
	Share	*Size*	*Diver*[1]	*Total*	*+ve*	*-ve*	*Broad*	*Medium*	*Detail*	*Diver*	*Size*	*Other*	*Total*	*+ve*	*-ve*	*Broad*	*Medium*	*Detail*
Class 1	2	3		5	42%	42%	50%	100%	33%	3	1	1	5	42%	8%	20%	100%	60%
Class 2	5	12	1	18	56%	25%	52%		75%	9	5	1	15	35%	19%	30%		50%
Class 3	5		3	8	50%	25%	14%	71%		13			13	68%	5%	44%	100%	
Class 4	8	3	1	12	63%	21%	58%	67%	100%	9	3	1	13	57%	0%	53%	33%	100%
Class 5	7	11	1	19	63%	13%	67%	63%	50%	9	12	2	23	70%	0%	76%	40%	100%
Total	*27*	*29*	*6*	*62*	*57%*	*23%*	53%	68%	56%	*43*	*21*	*5*	*69*	*53%*	*8%*	45%	64%	74%

Note.

[1] Diversity indicators were grouped with Other indicators

[2] Percentage of the total of indicators, which showed positive or negative results, the remaining percentage showed non-significant results.

[3] Percentage of the positive results per each industry classification level

[4] Classes of geographical units:

Class 1: state (US, Mexico), province (China), CSO region (UK), BEA area (US), region NUTS 2 (full coverage)

Class 2: county (UK), province (Italy, Spain), prefecture (Japan), department (France), COROP (Netherlands), region (Israel), CSO region (UK), region NUTS 3 (full coverage)

Class 3: labour zones: local labour systems (Italy), zones d'emploi (France), local labour market (Sweden) (full coverage)

Class 4: Zip-code (Netherlands, Spain), district LAU 1, region NUTS 4 (full coverage)

Class 5: SMA or MSA (USA), city, urban area (partial coverage)

Table 3-12: Number of positive results (dependent variables) per geographical unit

Geographical unit used[4]	***Number of dependent variables with positive results:***				***Total***
	MAR only	***Jacobs only***	***Both MAR and Jacobs***	***None of MAR and Jacobs***	
Class 1	4	4	1	3	12
Class 2	7	4	9	2	22
Class 3	2	6	4	2	14
Class 4	6	6	6		18
Class 5	6	6	11		23
Total	25	26	*31*	*7*	*89*

Note.

[4] See above

This part of analysis had as an initial objective to find out whether different kinds of externalities are associated with different geographical classes. The results in Table 3-11 seem to be quite balanced and do not show that the effects of Marshall or Jacobs externalities are influenced by the choice of geographical aggregation level. It seems, however, that the smaller the selected geographical unit is, the stronger are the effects encountered. Furthermore, with smaller geographical unit, there are more of Marshall and Jacobs simultaneous positive results and less of non significant or negative results. This is also observed by Glaeser *et al.* (1992), who notice that the magnitude of external effects increases as the geographical unit becomes smaller.

Table 3-11 also shows the allocation of positive variables into the categories of independent variables. Even though the general pattern (size-based and share-based indicators for Marshall externalities and diversity-based indicators for Jacobs) is still present, no consistent relationship between the size of the geographical unit and the number of positive independent variables in each category is observed.

3.4.4 Countries and regions

The differences in the impact of Marshall and Jacobs externalities on regional performance according to the country where the research is undertaken are studied. Table 3-13 and Table 3-14 group the studies according to the different countries examined and shows the positive results for both categories of externalities.

Table 3-13: Number of positive results per country by category of independent variable and industry classification level

Dependant variable	*MARSHALL EXTERNALITIES*									*JACOBS EXTERNALITIES*								
	Indicators with positive results:				*Total %*[2]		*Industry class.(+ve)*[3]			*Indicators with positive results:*				*Total %*[2]		*Industry class.(+ve)*[3]		
	Share	*Size*	*Diver*[1]	*Total*	*+ve*	*-ve*	*Broad*	*Medium*	*Detail*	*Diver*	*Size*	*Other*	*Total*	*+ve*	*-ve*	*Broad*	*Medium*	*Detail*
United States	*4*	*11*		*15*	56%	19%	73%	63%	44%	*8*	*5*	*2*	*15*	56%	4%	55%	44%	83%
United Kingdom	*1*	*10*		*11*	69%	19%	70%		67%	*4*	*1*	*1*	*6*	25%	25%	12%		57%
Italy	4	4	3	11	65%	29%	46%	83%		10	2		12	57%	14%	28%	100%	
Germany	1		1	2	50%	0%	50%			1		1	2	50%	0%	50%		
Spain	3	3		6	60%	10%	60%			3	6		9	75%	0%	75%		
Netherlands	5			5	38%	31%	30%		100%	6	2		8	62%	8%	50%		100%
France	1			1	20%	20%	0%	0%		4	0		4	57%	0%	40%	100%	
Finland	1		1	2	100%	0%	100%				1		1	33%	0%	0%		
Sweden			1	1	100%	0%							0	0%	0%			
Portugal	2			2	67%	33%		67%		1			1	33%	0%		33%	
Europe	2			2	67%	33%	50%		100%	2	1		3	75%	0%	50%		100%
Continental Europe total	*19*	*7*	*6*	*32*	55%	22%	*44%*	*70%*	*100%*	*27*	*13*	*1*	*41*	59%	6%	*46%*	*82%*	*100%*
Japan	2	1		3	75%	25%	75%			2	1		3	75%	0%	75%		
China				0	0%	100%	0%			1			1	50%	0%	50%		
Mexico	1			1	50%	50%	0%		100%		1	1	2	67%	0%			67%
Korea		1		1	50%	0%	0%	100%		1		1	2	67%	0%	100%	50%	
Brazil		1		1	100%	0%	0%						0	0%	0%	0%		
Israel		1		1	100%	0%					1		1	100%	0%			
Other total	2	*4*	*0*	*6*	55%	36%	*30%*	*100%*	*100%*	5	2	*1*	8	57%	0%	56%	*50%*	*67%*

Note.

[1] Diversity indicators were grouped with Other indicators

[2] Percentage of the total of indicators, which showed positive or negative results, the remaining percentage showed non-significant results.

[3] Percentage of the positive results per each industry classification level

Table 3-14: Number of positive results (dependent variables) per geographical region (country)

Geographical region	*Number of dependent variables with positive results:* *MAR only*	*Jacobs only*	*Both MAR and Jacobs*	*None of MAR and Jacobs*	*Total*
United States	7	8	*6*	*1*	*22*
United Kingdom	7	1	*3*		*11*
Italy	2	3	6	2	13
Germany	0	0	2		2
Spain	2	1	4		7
Netherlands	3	6	2	1	12
France	0	3	1	1	5
Finland	1	0	1		2
Sweden	1	0			1
Portugal	2	1			3
Europe	1	1	1		3
Continental Europe total	12	15	*17*	*4*	*48*
Japan	1	1	2		4
China	0	1		1	2
Mexico	0	0	1	1	2
Korea	0	1	1		2
Brazil	1	0			1
Israel	0	0	1		1
Other total	2	3	*5*	*2*	*12*

The results do not seem very different from one country to another. The exceptions in Table 3-13 are the United Kingdom, where the overwhelming majority of studies observed positive Marshall economies and to a certain extent the Netherlands, where Jacobs theory is mostly supported. Otherwise, the studies in all the other countries seem to find an even distribution of evidence for both specialization and diversity effects.

Some authors have carried out simultaneous studies of several countries and found quite comparable results, as Henderson (1986) for the US and Brazil. Nevertheless, other researchers have encountered distinct effects of the two externalities for different countries, as Beaudry and Breschi (2000, 2003) for the UK and Italy or Beaudry *et al.* (2001) for several European countries. In fact, the industrial and economic compositions of the studied countries differ and the spillover mechanisms actually may work quite differently. For example, Cingano and Schivardi (2004) note that Italy has quite a distinct productive system, which is characterized by areas with a substantial presence of small and medium enterprises in the traditional sectors, and which could be particularly conducive to interaction-induced externalities of the Marshall type. However, results found in this study do not show any substantial differences among countries.

The role of the local economic environment may vastly differ between Europe and the U.S. The often mentioned reasons are the different levels of labour mobility, which is much higher in the US and different unemployment rates that are higher in Europe. Both of these conditions could impact the spillover mechanism and influence the results. If the countries are grouped into the US, the UK, continental Europe and the rest (as in Table 3-13), some differences among these groups could indeed be seen, namely in the positive independent variable categories. In case of the US and the UK, positive results for Marshall variables are found mainly with size-based indicators, whereas for continental Europe they came usually from share-based indicators. This difference is probably not related to the various levels of labour mobility or different unemployment rates described above. In general, no systematic differences in the results caused by the choice of the European or the US data are found, the spillover mechanisms seem to be working in a similar fashion in both Europe and the US.

3.4.5 Period of observation

Another factor that may have influenced the results is the selected period of observation. Some studies survey the behaviour of the variables during prolonged periods of time, for example Boschma *et al.* (2005) cover around 130 years of industry development. During such an extended period, major events (like wars) might have had an enormous impact on the role of externalities and the definitions (of industries, regions, cities, etc.) might have changed substantially. Other studies, on the other hand, analyze the conditions and the relationships during as short period of time as one year (for example Costa-Campi and Viladecans-Marsal, 1999).

Moreover, even if the time range is of comparable length, it may still matter that the period is not exactly the same. Externalities will have stronger impact during economically dynamic time periods and the results then cannot be comparable with their effects during the periods of the relative economic stagnation (Glaeser *et al.*, 1992). Combes (2000a) conjectures that depending on the economic cycles there may be asymmetric effects associated to specialisation: Marshall economies would enhance local growth during expansion periods, but it would also favour employment decline during recessions due to inflexibilities and rigidities. This hypothesis, however, calls for further testing.

3.4.6 Size of the firms

The last factor to be briefly mentioned here is the effect of firm size on the role of externalities in regional performance. Only few of the reviewed articles distinguished between the firms of different sizes and these are the firm-level studies. The studies that did, however, are in agreement: Marshall economies have a positive or more profound impact on small (or non-corporate) firms (Beardsell and Henderson, 1999; Mukkala, 2004; Van der Panne, 2004), whereas Jacobs economies are more advantageous for large (or corporate) firms (Capello, 2002; Henderson, 2003). Acs and Audrtesch (1988, 1990) study innovative intensity and show that small firms are more innovative in proportion than large firms even though the latter introduced a greater number of product innovations.

3.5 Conclusions

The reviewed empirical work has provided substantial academic support for the positive impact of both Marshall and Jacobs externalities on regional performance. As for their negative effects, the empirical evidence shows that specialization of a region may hinder economic growth, whereas diversification is much less likely to produce this negative impact. The results of these studies are, however, often conflicting and mutually contradictory. This can be explained by differences in the strength of agglomeration forces across industries, countries or time periods, but also by some methodological issues.

The most obvious differences among the studies are the ones associated with the choice of independent and dependent variables. The common indicators for the former are specialization and diversity. The two most frequently used measures of specialization, location quotient and own industry employment, did not produce balanced effects in the regressions. A greater number of studies find negative results for Marshall externalities when using the location quotient than when using own industry employment, whereas the chance of observing a positive impact of specialization is similar in both cases. Furthermore, the choice of diversity measure seems to be critical to the observation of the presence of Jacobs externalities. While the use of other industry employment would probably result in negative or no effects of diversified region, Gini index of diversity would provide positive results. With the selection of Hirschman-Herfindahl index, there is an equal number of studies that find positive or neutral results on Jacobs economies.

There seem to be distinct effects of each of the externalities on the different performance measures, used as dependent variables. It is shown that Jacobs externalities favour economic growth more than do Marshall economies. In contrast, if the influence of the industrial composition on productivity growth is studied, Marshall's theory is more often supported. Only the studies using the dependent variable for assessing the impact on innovative activity have provided balanced support for both theories. Furthermore, the results differed according to the level, at which the dependent variable is analyzed. The results show that a slightly greater number of studies find evidence of Jacobs externalities at the firm level while at the regional level, Marshall's thesis is dominant (especially if the impact on economic growth is studied). It is also shown that a relative size of an industry (its share in the region) has a stronger positive impact in the form of Marshall externalities on economic performance at the regional level, while it is the absolute size of industry which matters more for the growth of individual firms in a region. As for Jacobs externalities, it is mainly the industrial base diversity, and to a much lesser extent the regional market size, that promotes the regional growth.

The suitability of the specialized or diversified regions for particular industries is assessed in a more detailed analysis of industrial sectors. Although not overwhelming, there is some evidence that in low tech sectors, Marshall externalities have stronger effects than Jacobs externalities. The situation in medium tech sectors yields similar results for both theories, but differs for the high tech sectors. The latter slightly favour diversified regions, while the effects of Marshall externalities are less pronounced. Diversification also appears to be a growth promoter in services. Both relative and absolute sizes of the given industry signify the presence of Marshall externalities for low and medium tech industries, while in high tech sectors it is mainly the absolute size of the industry that matters. In the case of Jacobs externalities, for all industrial types it is uniquely the industrial base diversity, and not the size of the local market, that promotes regional growth. Moreover, it is shown that the role of externalities varies according to the maturity the industry. Jacobs externalities predominate in the early stages of the industry life cycle, whereas Marshall externalities enter at a later point, and in the end specialization will in fact hinder economic growth.

Another factor which was examined is the level of industrial classification used for the analysis. An analysis of the results per industry class has shown that MAR effects are slightly more prone to show up at the broad level of detail, the probabilities of detecting MAR or Jacobs effects are quite comparable at the medium level and the Jacobs effects will decidedly appear more often when the detailed classification is used. The 3-digit classification was thus suggested to be a threshold, at which specialization and diversity are less distinguishable from one another, before which it is specialization and beyond which it is diversity that are favoured.

The geographical dimension is evaluated from two points of view: the level of geographical aggregation and the choice of the region or country. The results show that the relative effects of the Marshall or Jacobs externalities are not influenced by the choice of geographical aggregation level. It seems, however, that the smaller the selected geographical unit the more positive effects are in general encountered. No country results seem to favour one theory above the other, and the spillover mechanisms are working in a similar fashion in both Europe and the US.

Finally, several studies seem to be in an agreement that the Marshall economies have a positive or more profound impact on small firms, whereas the Jacobs economies are more advantageous for large companies.

There are quite important implications of this investigation for public policy. Whether the externalities needed for a successful development of a particular industry and a particular region are of Marshall or Jacobs kind may affect the design of a regional development strategy. This chapter suggests that in regions with mature, low tech industries, regional policy should emphasize the development of a narrow set of economic activities in the region, which will presumably lead to greater productivity. In high tech regions, on the other hand, policy should focus on the creation of a diverse set of economic activities, which should enhance economic development. However, given such contrasting opinions and conflicting conclusions, any regional development policy which selects, supports or discriminates certain industrial activities or technologies should be applied with caution until the issue is fully clarified. Much more research is needed to fully understand such an abstract phenomenon as knowledge spillovers, their localized character and their impact on the innovative process and regional performance.

CHAPTER 4

INNOVATION IN CANADIAN BIOTECHNOLOGY CLUSTERS

Canada has a small population dispersed over a large geographical area and its private sector is dominated by small-sized and medium-sized companies. As a consequence, research and development has to concentrate in geographical agglomerations and clusters in order to contribute to an efficient innovation system. The context of this chapter is the biotechnology field, which should presumably benefit from the types of knowledge spillovers and information exchanges that are facilitated by spatial clustering. The knowledge base in biotechnology is largely tacit and uncodifiable, which are generally favourable conditions for knowledge spillovers in agglomeration economies. Niosi and Bas (2001) find that biotechnology activity in Canada is indeed clustered and is mainly concentrated in three large cities – Toronto, Montreal and Vancouver, where most patents and venture capital are located. Niosi and Bas note that clusters have also developed around medium-sized urban agglomerations, such as Ottawa, Edmonton, and Calgary, or specialized clusters around some smaller cities. They also argue that it is the population of the metropolitan area and the local university research, which are key factors explaining the size, location and characteristics of these clusters. They identify universities, government laboratories and a few large firms as the main anchor tenants in Canadian biotechnology clusters. Aharonson *et al.* (2004) argue that, in Canada, clustered biotechnology companies are eight times more innovative than the ones that are remotely located. The largest effects were observed for firms located in clusters strong in their own specialization. Niosi and Banik (200) also find that biotechnology companies in the clusters of Montreal, Toronto and Vancouver perform better than companies outside these clusters.

Another line of research aims at shedding light on the determinants of differential growth in biotechnology companies in Canada. It is often argued that alliances and cooperation are an indispensable element in the success of small firms. Niosi (2003) suggests that international alliances with large pharmaceutical corporations are the main determinant of growth in Canadian biotechnology and that timely alliances are also the critical key factor for the survival of the new biotech firms. Oliver (1994) empirically confirms that this inability of a new biotechnology company to form inter-firm alliances is associated with organizational death. Niosi (2003) however argues that the success at forging suitable alliances alone does not sufficiently explain differential growth in biotechnology companies. He adds that the quantity of patents, the amount of venture capital, the size of exports and the specialization in human health products play an extremely important role as well. Queenton and Niosi (2003) propose two other determinants of rapid growth: the quality of patents and the presence of star scientists in biotechnology firms. According to them, Canadian biotechnology clusters are strongly related to high-class academic research and star scientists working in universities. Their study also highlights the importance of geographical proximity of star scientists for obtaining venture capital, and for starting and growing biotechnology firms. It was also confirmed that in Canada many of the star scientists capitalise on their knowledge through firm start-ups. One third of Canadian biotechnology firms are estimated to be university spin-offs (Niosi, 2003).

In summary, growth and continued health of Canadian biotechnology clusters are among others dependent upon the presence of major attractors such as research universities and governmental laboratories active in biotechnology, innovative propensity of the local scientists (*i.e.*, the existence of star scientists), formation of alliances and active cooperation among biotechnology firms, composition of biotechnology fields in the cluster (*i.e.*, the focus on the health-related products) and presence of the largely innovative biotechnology firms (with patents of a high quality and quantity). This chapter intends to address most of the above factors. It aims at understanding the creation of innovation in Canadian biotechnology clusters and its main objective is to identify, analyze and describe Canadian biotechnology clusters based on the characterization of the quality and quantity of their innovative outputs, the nature of biotechnology activities which are carried out in these clusters, the characteristics of the patent-owning entities and their propensity to collaborate. The major contribution lies in embracing a cluster approach, *i.e.*, all analyses are made at the cluster level (all properties and characteristics are calculated per cluster). Most of the studies providing a descriptive profile of innovation in Canadian biotechnology are carried out and presented at either the province or firm level. Moreover, a full picture of innovation in Canadian biotechnology is built by including all of the biotechnology agglomerations in Canada. Previous cluster-based studies have focused mainly on two or three major Canadian biotechnology clusters, but little is known about the smaller concentrations, which are less active in biotechnology. Finally, this research is based on the complete database of all the Canadian biotechnology patents registered with the United States Patents and Trademarks Office (USPTO), which no other study has done so far. Previous researchers usually adopted the approach of analyzing only representative samples.

4.1 Canadian biotechnology clusters

A cluster in this thesis is defined as a geographically continuous region active in biotechnology (as measured by patent production). A summary of the basic statistics regarding the 12 identified clusters defined in such a way is presented in Table 4-1. Also, see Figure 6-2 in Chapter 6 which graphically demonstrates the number of patents and inventors in each cluster.

The Toronto cluster decisively leads in the number of Canadian biotechnology patents; it has almost twice the number of patents of Montreal, which in turn has almost twice as many patents as Vancouver. Even Ottawa has more patents than an important biotechnology cluster such as Vancouver. Most of the biotechnology activity carried out in Canada takes place inside these clusters, usually the few main ones. Only 2% of the patents have no assignee[17] inside these clusters. These are found most often in Ontario (19 patents) or Alberta (13 patents). There are very few patents (1%) with co-assignees from multiple Canadian clusters. The lack of common inter-cluster ownership of patents suggests that there is very little cooperation at the assignee level between clusters and if there is, ownership of patents is not shared (for more details on cooperation among the institutions see section 4.4. of this chapter, for the cooperation among the inventors see Chapters 5 and 6).

Table 4-1: Summary of the results for biotechnology clusters

Biotechnology cluster	*Number of patents**	*as % of all patents*	*Claims (average)*	*Number of inventors***	*as % of all inventors*	*Patents per inventor****
Toronto	834	34%	14.6	927	29%	1.44
Montreal	466	19%	14.7	698	22%	1.05
Vancouver	255	10%	19.9	411	13%	0.95
Edmonton	153	6%	13.1	210	7%	1.21
Calgary	127	5%	16.8	91	3%	2.19
Saskatoon	98	4%	20.1	147	5%	1.04
Winnipeg	33	1%	13.8	77	2%	0.91
Kingston	63	3%	16.7	94	3%	1.01
Ottawa	279	11%	15.3	224	7%	1.26
Quebec	57	2%	15.8	127	4%	0.97
Halifax	20	1%	16.2	33	1%	1.06
Sherbrooke	16	1%	13.2	26	1%	1.07
outside clusters	47	2%	19.4	159	5%	1.16
co-assignees from multiple clusters	37	1%	12.8	-	-	-
CANADA	**Σ 2485**	100%	15.6	**Σ 3224**	100%	1.21

* The numbers are based on the residence of assignees and only the patents with at least one Canadian assignee are thus included (i.e., 753 foreign-assigned patents and 312 non-assigned patents are excluded).

** Inventors with multiple addresses (who patented while living in several clusters) were assigned to only one cluster.

*** Patents per inventor are counted as the number of patents co-invented by at least one inventor from the cluster divided by the number of inventors who at least once patented while living in that cluster.

---

[17] Patent assignee is an entity (original or legal company, organization or person) that is registered as proprietor of the patent or patent application.

Assuming that a greater number of claims[18] corresponds to the higher value/quality of a patent, Table 4-1 shows that the quality of the patents whose assignee is from the Vancouver or Saskatoon clusters is much higher than the quality of other patents. In addition, the patents whose owners reside outside the clusters are observed to have a much higher quality on average. Moreover, note that in Table 4-7 Canadian-owned patents have lower than average quality, while patents owned by foreign assignees have higher than average quality. American-owned patents in particular have higher quality than the ones owned solely by Canadians. These results are in agreement with Tong and Frame (1994) who have compared the random sample of over 7000 patents in five countries (Japan, France, W. Germany, the UK and the US) and found that the US has the largest average number of claims per patent (13.8).

4.2 Patent ownership in Canadian biotechnology clusters

In order to understand the institutional composition of the biotechnology clusters the ownership of the patents has been examined. The patents were divided according to the nature of the entity to which they were assigned. Table 4-2 shows the distribution of the patents based on the category of the patent owner.

Around half of the patents are assigned solely to companies, much less to universities or to governmental institutions. Biotechnology is a scientific field with potentially high financial revenues, which probably explains the high entrepreneurial interest and consequently the high representation of the private sphere among the biotechnology assignees. Commercial interests push the private biotechnology

---

[18] Patent claims are a series of numbered expressions describing the invention in technical terms and defining the extent of the protection conferred by a patent (the legal scope of the patent). A high number of patent claims is an indication that an innovation is broader and has a greater potential profitability. It has been frequently suggested and empirically demonstrated (see for example Tong and Frame, 1994) that the number of claims is significantly and consistently indicative of higher value patents. The conclusions of most of the papers on patent value reviewed by van Zeebroeck and van Pottelsberghe de la Potterie (2006) are supportive of the positive association of the number of claims with patent value. Lanjouw and Schankerman (2004) have suggested that specifically in the biotechnology field the number of claims is the most important indicator of patent quality.

companies to strictly protect their most important assets, intellectual property, by employing appropriate mechanisms (such as patents), whereas the registration of the university or governmental inventions at the patenting offices may not seem so crucial to the individual inventors, who themselves may not particularly care about the financial well-being of the institution. Moreover, the main objective of a university or a research lab is not to make money (in comparison with the private company) and the process of intellectual property protection and invention commercialization thus may not be given as high importance.

Table 4-2: Patents by category of assignees

Assignees' category	*Number of patents*	*As % of all patents*	*Claims (average)*	*Dominant cluster*
firm (single or multiple)	1792	50%	17.0	Toronto
university (single or multiple)	743	21%	15.7	Montreal
government* (single or multiple)	338	9%	13.6	Ottawa
hospital (single or multiple)	137	4%	13.5	Toronto
firm-university	67	2%	17.1	Toronto, Montreal
individual (single or multiple)	25	1%	11.4	Edmonton
government-university	32	1%	16.2	Ottawa
hospital-university	33	1%	13.6	Toronto
other categories**	71	2%	17.38	-
non-assigned	312	9%	14.4	-
TOTAL	**Σ 3550**	100%	15.98	-

* Government assignees include all the federal or provincial laboratories and research institutions, Canadian ministries and ministers, Her Majesty the Queen in right of Canada, etc.

**Includes the following co-assignees categories: firm-government, firm-hospital, firm-hospital-university, firm-government-university, government-hospital, firm-individual, hospital-individual, government-individual, individual-university, hospital-government, hospital-government-university and individual-firm-hospital-government

Canada has the second highest share of industry-financed research in the academic sector among the G7 countries (Germany has the highest score). Industry financed over 8% of Canadian university R&D activities in 2005 (OECD, 2007). This is suggestive of strong linkages between industry and universities, which were however not observed in the data. In an ideal world, frequent cooperation between firms and universities should be revealed by a higher number of co-assignments of biotechnology patents. Only 67 patents were co-assigned to a firm and a university simultaneously, which represents only 2% of all patents. The particular contractual arrangements regarding intellectual property rights between universities and firms are probably the reason why the patents resulting from joint research projects are assigned uniquely to the firm or to the university.

Furthermore, patents whose assignees or co-assignees are firms have higher value (as measured by the average number of claims for patents in each category) than most of the ones whose owners are not companies. The patents of the lowest value are generally owned by individuals, usually the inventors themselves. This may suggest that the reason why an individual researcher did not offer the patent to any company or did not pursue the commercialization by himself (by founding the biotech firm) is that he probably did not perceive the patent to be of a high quality and predicted that the chances for lucrative commercialization were slim.

Figure 4-1 shows the proportions of patents assigned to different entities in the most common categories. The Toronto cluster possesses the largest (75%) proportion of patents assigned to firms. The portion of the Montreal company-owned patents is considerably lower (58%). Industrial biotechnology research is highly concentrated in the big clusters such as Toronto or Montreal, while university research is spread over the small Canadian clusters. The enormous share of the patents in Ottawa assigned to government entities strikes at first sight. Ottawa, as the capital of Canada, hosts many federal government research institutions producing biotechnology patents. A certain number of patents which do not involve any local inventive element and are generated outside the Ottawa cluster, are still assigned to, or being represented by, the federal institutions in Ottawa, for example the National Research Council, various ministers or Her Majesty the Queen (who herself is an owner of 92 biotechnology patents and is hence a biggest individual biotechnology patent owner in Canada!). The National Research Council of Canada has five national biotechnology institutes throughout the country (Montreal, Vancouver, Ottawa, Saskatoon and Halifax) but 99% of the patents owned by the National Research Council are assigned to its central office in Ottawa.

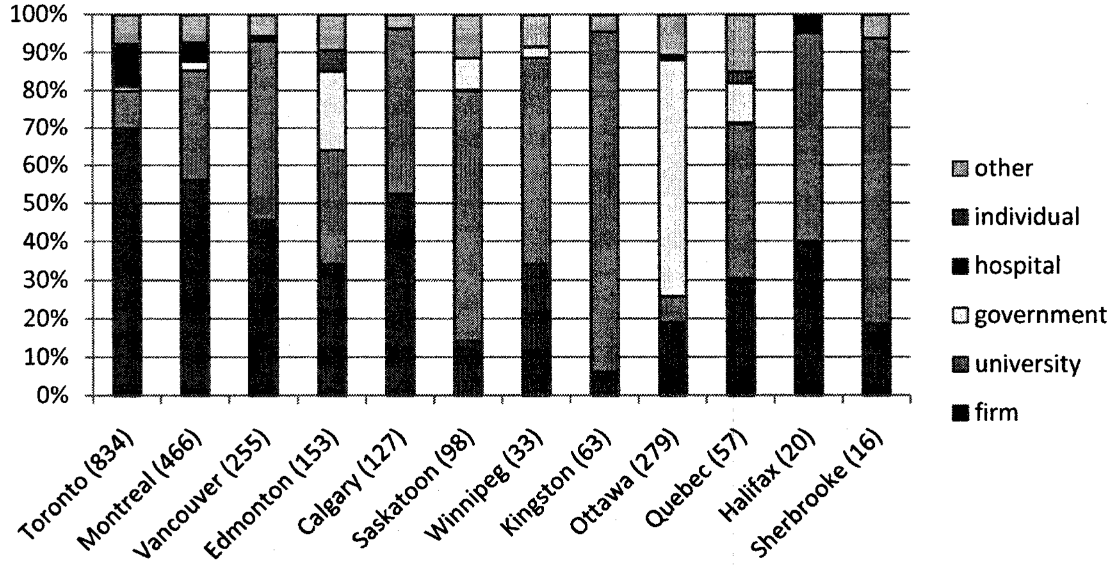


Figure 4-1: Shares of patents assigned to the various entities in each cluster (the number in parentheses shows the total number of patents in the cluster)

The picture is also blurred by the fact that in some clusters the patents produced by university hospitals or hospital research centres affiliated to universities may have been assigned to the hospitals themselves, while in other clusters they are assigned directly to universities. This probably explains the very high percentage of university patents found in clusters in which at the same time there is not a single hospital-assigned patent (as in Vancouver, Calgary, Saskatoon, Kingston, Winnipeg and Sherbrooke), while in Toronto, there are more patents actually assigned to hospitals than to universities. Even if this fact is taken into consideration, the shares of patents assigned to universities in these five clusters are still substantial, whereas the portions of the university-owned patents in the Toronto and Ottawa clusters are alarmingly low. The differences in university patenting among the clusters are probably related to the distinct intellectual property (IP) rules and policies which, in Canada, are governed by the universities themselves. In 2003, 78% of Canadian universities actively participated in managing intellectual property, but formal requirements to disclose inventions existed only in 45% of universities (Read, 2005).

The rules regarding the ownership of the IP rights at the universities with the highest numbers of biotechnology patents (see Table 4-3) were investigated. In the case where the IP rights are owned by an inventor or jointly by both an inventor and a university at the time of invention creation, the inventor usually has the option to either commercialize the invention himself or assign the IP rights to the university, where a technology transfer office will take care of the commercialization process. In many cases the inventions are owned by default by the university who decides whether to commercialize the invention or not. Table 4-3 presents the distribution of the net revenue based on whether the ownership of the invention is retained by the inventor or by the university. This shows that inventors at various universities have quite diverse opportunities and motivations for the commercialization of their inventions. An academic inventor who retains the IP rights may consider the patenting of his invention as an expensive, lengthy, risky, drudging and usually not particularly profitable process. Furthermore, as publication and not patenting is more rewarding in one's academic carrier, a prolonged patent application process can delay the inventor's ability to publish.[19] The university technology transfer office usually seeks to commercialize the IP more actively, efficiently and professionally than if the commercialization is left to the individual academic. Therefore, a university which reserves exclusive IP rights to all the university-generated inventions and/or has well functioning technology transfer office will usually show higher patent counts. In contrast, much lower are the number of patents owned by universities where the academics may retain the IP rights to their inventions themselves and where the transfer offices do not manage to offer attractive commercialization alternatives to the inventors.

---

[19] In Canada, the USA and Mexico an inventor has one year after publishing his invention to file a patent application. Nevertheless, in most of the other countries, the novelty of the invention is destroyed by publishing an enabling description of the invention before filing for a patent protection.

Table 4-3: IP ownership and the distribution of the net revenue at the most prolific universities patenting in biotechnology in Canada

			IP rights ownership retained by and/or invention commercialized by			
			THE INVENTOR		*THE UNIVERSITY**	
			Distribution of the net revenue:		*Distribution of the net revenue*	
	Number of biotechnology patents	*IP rights ownership at the time of invention*	*Inventor*	*University*	*Inventor*	*University*
McGill University	123	Joint	first $100 000: 80%, then 70%	first $100 000: 20%, then 30%	first $10 000: 100%, then 60%	first $10 000: 0%, then 40%
U. of British Columbia	114	University			50%	50%
University of Saskatchewan	78	University			50%	50%
Queen's University	61	Inventor	first $500 000: 100%, then 75%	first $500 000: 0%, then 25%	*negotiated*	*negotiated*
University of Calgary	59	Inventor	75%-90%	25%-10%	50%	50%
University of Alberta	57	Inventor	66.6%	33.3%	33.3%	66.6%
University of Guelph	35	University			first $100 000: 75%, then 25%	first $100 000: 25%, then 75%
Université de Montreal	32	Inventor	*negotiated*	*negotiated*	*negotiated*	*negotiated*
Université Laval	32	University			50%	50%
University of Toronto	28	Joint	75%	25%	60%	40%
University of Ottawa	26	University			first $100 000: 80%, then 50%	first $100 000: 20%, then 50%
University of Manitoba	22	Joint	*negotiated*	*negotiated*	*negotiated*	*negotiated*
University of Victoria	20	Inventor	if <$5,000/year: 100%, else 80%	if <$5,000/year: 0%, else 20%	*negotiated*	*negotiated*

* Some universities have founded special non-for-profit organizations in order to commercialize the IP of the university-generated research. These organizations may have exclusive rights to the university IP. However, it is not distinguished here whether the invention is owned and/or commercialized directly by the university or by this organization.

Table 4-4 shows the main statistics concerning the assignees of the biotechnology patents whose inventors include at least one Canadian inventor. The first column includes all categories of assignees, while the second one only counts the assignees recognized as private firms. Comparing the current results with external resources, it was found that according to Statistics Canada (2007) the total count of innovative biotechnology firms (i.e., firms that are engaged in the development of new products or processes) was 542, which seems considerably more than the obtained count of 299. The used methodology obviously underestimates the number of biotechnology firms, because only the companies which have at least one biotechnology patent registered at the USPTO are considered. Firms may have been left uncounted for various reasons: First, this method excludes all the biotech companies that patent solely at different patent offices (e.g. CIPO, EPO). The number of such companies is unknown, however, since Canadian inventors usually do not patent solely in Canada (as explained earlier) it is assumed that the number of patents registered exclusively at the CIPO are not substantial. Some of the inventors (particularly the ones who collaborate with European researchers) may nevertheless have chosen to file their patent application both at the CIPO and the EPO. Second, it obviously also excludes the firms which do not patent any inventions at all. Due to the high patenting propensity of biotechnology firms it can be presumed that biotechnology companies would rarely choose not to patent at all. The main focus of this research is on innovation and thus the exclusion of a company which does not pursue any patentable innovative activity does not change the picture significantly. Third, an innovative biotechnology firm will not be included if it prefers (probably for strategic reasons) an alternative means of intellectual property protection such as technology transfer agreements or licensing. Comparison with Niosi (2005) offers even more distinct findings: Niosi presents both the count of all Canadian biotechnology firms and the count of Canadian biotechnology firms with patents (issued by the USPTO). He suggests that it is Montreal that leads in both of these categories, which is quite contrary to the results obtained. His database is however fairly limited, since it includes only 24 firms with patents in Montreal and 22 of them in Toronto.

Table 4-4: Results regarding assignees as counted per biotechnology cluster

Biotechnology cluster	*Number of assignees*	*Number of firms*[*]	*Patents per assignee*	*Patents per firm*[**]	*Inventors per assignee*[***]
Toronto	144	110	5.79	5.38	6.44
Montreal	77	55	6.05	4.80	9.06
Vancouver	51	44	5.00	2.70	8.06
Edmonton	29	19	5.28	2.89	7.24
Calgary	13	11	9.77	6.18	7.00
Saskatoon	9	6	10.89	2.50	16.33
Winnipeg	9	7	3.67	1.71	8.56
Kingston	3	2	21.00	2.00	31.33
Ottawa	28	11	9.96	5.00	8.00
Quebec	18	11	3.17	1.82	7.06
Halifax	5	3	4.00	2.67	6.60
Sherbrooke	4	2	4.00	1.50	6.50
outside clusters	28	18	1.68	1.33	5.68
CANADA	**Σ 418**	**Σ 299**	5.86	4.10	7.71

[*] Only inventive firms (i.e. those which have produced at least one biotechnology patent) are counted.
[**] Number of patents assigned to firms divided by the number of inventive firms in the cluster.
[***] Number of all inventors divided by the number of assignees in the cluster.

It is interesting to note in Table 4-4 that the lowest number of assignees (3 assignees) is found in the Kingston cluster, which by no means counts among the smallest biotech clusters with 63 patents and 94 inventors. An overwhelming majority of the Kingston's patents are produced at Queen's University and there are in fact only two innovative companies in the cluster. A similar situation exists in Saskatoon with the University of Saskatchewan producing or co-producing almost 73% of all the patents. Saskatoon also hosts the NRC Plant Biotechnology institute which may generate a large portion of its patents assigned to the Ottawa NRC head quarters. The "patent per assignee" ratios in the next column are thus considerably higher for these two clusters and the "patents per firm" ratios are understandably much lower. Toronto and Montreal show quite comparable numbers of patents whether measured per assignee or per firm. However, the number of patents produced on average by firms in the Vancouver cluster is considerably lower than the one of the other two major clusters. While a private company in Toronto or Montreal has registered on average around 5 biotechnology patents at the USPTO, the Vancouver (or Edmonton) companies have been granted only 2-3 patents. As for the number of inventors per institution, it is especially high in Montreal. Calgary and Ottawa are clusters with relatively high numbers of patents, which even exceed the number of inventors. Their ratios of patents per assignee or patent per firm are consequently also fairly high. This probably means that the institutions in these clusters involve a large number of biotechnology researchers (e.g. in Ottawa, NRC and Her Majesty the Queen are ranked fourth and fifth as assignees with the greatest number of patents), but also the biotech companies are probably larger (as the high ratio of "inventors per assignee" suggests). Otherwise, it could be generally stated that smaller clusters have a lower number of patents per institution or per firm, implying either that companies in these clusters are on average smaller as well or that they are simply patenting less.

Canadian assignees which are the full or partial owners of more than 20 biotechnology patents at the USPTO are listed in Table 4-5. In addition to the information on the number of patents, the number of papers in biotechnology is also shown (provided by the National Research Council Canada, 2005) for the institutions most active in biotechnology. It has already been mentioned that the National Research Council of Canada possesses five biotechnology related institutes throughout Canada; unfortunately, the assignee is more often than not the main office in Ottawa and as such does not allow a regional distinction.

Table 4-5: Assignees with Canadian residence with 20 or more patents filed with the USPTO

	Assignee	*Number of patents*	*Number of papers*[*]	*Cluster*	*Province*	*Assignee's category*
1	McGill University	123	372	Montreal	QC	university
2	Connaught Laboratories Ltd	118		Toronto	ON	firm
3	University of British Columbia	114	308	Vancouver	BC	university
4	National Research Council of Canada	95	160	Ottawa	ON	governmen
5	Her Majesty the Queen of Canada	92		Ottawa	ON	governmen
6	University of Saskatchewan	78	170	Saskatoon	SK	university
7	Hospital for Sick Children	71	109	Toronto	ON	hospital
8	Aventis Pasteur Ltd	63	13	Toronto	ON	firm
9	Queen's University	61		Kingston	ON	university
10	University of Calgary	59	189	Calgary	AB	university
11	University of Alberta	57	244	Edmonton	AB	university
12	Allelix Biopharmaceutical	52		Toronto	ON	firm
13	Merck Frosst Canada Inc.	42	14	Montreal	QC	firm
14	Visible Genetics Inc	40		Toronto	ON	firm
15	University of Guelph	35	223	Toronto	ON	university
16	Alberta Research Council	34		Edmonton	AB	governmen
17	Université de Montreal	32	209	Montreal	QC	university
18	Université Laval	32	205	Quebec	QC	university
19	Syn X Pharma	32		Toronto	ON	firm
20	Mount Sinai Hospital	31	108	Toronto	ON	hospital
21	University of Toronto	28	533	Toronto	ON	university
22	University of Ottawa	26		Ottawa	ON	university
23	Canadian Patents and Development	26		Ottawa	ON	governmen
24	Boehringer Ingelheim Canada Ltd	26		Montreal	QC	firm
25	Adherex	25		Ottawa	ON	firm
26	University of Manitoba	22		Winnipeg	MB	university
27	NPS Allelix Corp.	22		Toronto	ON	firm
28	Spectral Diagnostics Inc.	21		Toronto	ON	firm
29	Ontario Cancer Institute	21		Toronto	ON	Hospital
30	University of Victoria	20		Vancouver	BC	university

[*] Source: National Research Council (2005). Information is provided only where available.

Table 4-5 confirms that biotechnology innovation is strongly based on publicly-funded research. Out of the first thirty assignees with the highest number of biotechnology patents there are 13 universities, 5 government institutions and 2 hospitals. The most important producers of patents are universities with McGill University (123 patents) heading the league table. Universities are also unsurprisingly the most active institutions in terms of the scientific papers production. Here the apparent leader is the University of Toronto (533 papers), which however owns a rather low number of patents (28 patents) in comparison. This shows again that in spite of the high quality research which is conducted at University of Toronto, not many inventions have probably reached the hands of the university technology transfer offices. During the last 20 years, the intellectual property policies at the University of Toronto did not encourage the inventors to assign the patents to the University. Moreover, in many cases, even though the inventors fully owned the IP rights, the University of Toronto was still engaged in commercialization of their inventions. However, the university has recently made many changes into its IP policy, and it remains to be seen in the coming years how these changes will be reflected in the number of university-assigned patents[20]. Other universities with a disproportionately higher publication record (in comparison with the number of patents) are Université de Montreal and Université Laval, which both have over 200 papers but only 32 patents.

The contribution of government institutions to the biotechnology research and development is also substantial: among the five highest ranking patent holders is the National Research Council of Canada[21] (95 patents), the Government of Canada's premier agency for research and development and Her Majesty the Queen in right of Canada (92 patents) usually representing various federal ministries (agriculture, health, national defence). The Alberta Research Council, a research agency owned by the province of Alberta, holds 34 patents and the Canadian Patents and Development, the agency which was engaged (before it was disbanded in the late 1980's) in commercializing the research performed at government labs, possesses 26 patents. The most active government institutions in biotechnology research are the National Research Council (160 papers) and Agriculture and Agri-Food Canada (191 papers), which is however the owner of only 8 patents.

---

[20] Information gathered during the conversation with the technology transfer office at University of Toronto

[21] Inventors' addresses will indicate where the research has actually taken pace.

The Hospital for Sick Children (71 patents and 109 papers) and the Mount Sinai Hospital (31 patents and 108 papers), which are both affiliated to the University of Toronto, lead the patent league among hospitals. These could explain how patents "escape" from the ownership of the University of Toronto. According to the university's IP policies, the patents are usually assigned to the institution where the research takes place physically. The university professor who is at the same time a doctor at one of the university-affiliated hospitals will probably carry out most of his research at the hospital, which will thereby become the patent owner[22].

Finally, a number of private companies are also the owners of a considerable number of biotechnology patents. The most inventive firms reside mainly in Toronto (e.g. Connaught Laboratories with 118 patents, Aventis Pasteur with 63 patents, Allelix Biopharmaceutical with 52 patents), but also in Montreal (Merck Frosst Canada with 42 patents, Boehringer Ingelheim Canada with 26 patents) or in Ottawa (Adherex with 25 patents). Only the patents registered under the Canadian residence of an assignee are counted (this excludes subsidiaries with the same name but with an address outside Canada). As expected, the number of papers published by private companies is relatively small (the highest is 14 papers by Merck Frosst Canada and 13 papers by Aventis Pasteur), since they prefer to protect their assets by patenting rather than revealing them into public domain through scientific papers.

4.3 Biotechnology field specialization in clusters

As a next step, the various biotechnology fields and their representation in the database were investigated. Various biotechnology fields were grouped according to the final use of the products into four categories: health-related biotechnology, agriculture and food related biotechnology, environmental and industrial biotech, and other

---

[22] Information gathered during the conversation with the technology transfer office at University of Toronto

biotechnology[23]. It was found that health-related biotechnology clearly represents the greatest proportion (78%) of all patents in the database, while agriculture and food related biotechnology accounts for 10% and environmental and industrial biotech only for 5% of all the patents. Table 4-6 includes for comparison the allocation of the biotechnology firms into these categories by different measures provided by Statistics Canada (2007). The table shows that the proportions of the health-related biotech patents (78%), the profits in the health biotech field (70%) as well as health-related R&D expenditures (87%) are all considerably higher than the proportion of biotechnology companies belonging to the health-related biotechnology (56%). This likely reflects the distinct characteristics of entrepreneurship in the health biotechnology. A company in this field would often have significant R&D expenditures and experience long development times before it has many products on the market. Afterwards, however, it would reap high profits, often far exceeding those of the firms in other biotech fields. Therefore it was found the health-related companies having on average much higher proportions of R&D expenditures, but even higher shares of the biotechnology profits than in the agricultural, environmental or industrial biotechnology fields. Note that the allocation of the biotechnology activity in Canada into the four main specializations according to the R&D expenditures roughly corresponds to the obtained results based on

---

[23] The definitions of biotech fields used in this thesis are as follows:

Health-related: human or animal health - pharmaceuticals, diagnostics, therapeutics), bioinformatics (gene sequencing, peptide or protein sequence, genomics, gene expressions etc.), nanobiotechnology, devices and apparatus specific for the use in health-related biotech;

Agriculture and food related: plant based agriculture (including fertilizers, manure, composting, herbicides and insecticides, etc.), food and edible materials for humans, feeding compositions for animals, nutrition (but not with specific therapeutic uses or vitamins, etc., which belong to the health-related biotech);

Environmental and industrial: environmental (biofuels, bioremediation, biodegradation, reutilization or destruction of garbage and waste, bioleaching etc.), industrial biotech (processing of metals, production of chemicals, other manufacturing processes, etc.), bioprocess technology (biocatalysis, bioseparation, biofilter, bioreactor, etc.);

Other - multiple uses in more than one of the above categories, non-specific biotech lab equipment (devices, apparatus, etc.) or completely other uses (for ex. fingertips in police investigation) the number of patents. Both of these indicators are closely related to biotechnology innovation: R&D expenditures measure the innovative process input, while the number of patents is usually considered to be an indicator of innovative output. A strong relationship between patent numbers and R&D expenditures has been also observed by Griliches (1990) who even suggested using the patents as an indicator of inventive input.

Table 4-6: The number of the biotechnology patents and firms by the biotech field

Biotech field	*Biotechnology patents*		*Biotechnology firms**		*Biotechnology revenues**		*R&D expenditures*	
	number	*%*	*number*	*%*	*million $*	*%*	*million $*	*%*
Health	2777	78%	303	56%	2955	70%	1486	87%
Agriculture-food	343	10%	140	26%	1075	26%	157	9%
Environment-industry	168	5%	54	10%	121	3%	34	2%
Other	262	7%	45	8%	41	1%	27	2%
All	3550	100%	542	100%	4192	100%	1704	100

*Source: Statistics Canada (2007)

The proportions of the patents of each biotech specialization as granted to assignees in the various categories are shown in Figure 4-2 which confirms that the major share of patents for all kinds of assignees pertains to health-related biotechnology. Not surprisingly, hospitals and other health institutions have a complete focus (100%) on health-related biotechnology. The highest proportion (93%) of the health-related biotech research is carried out (after hospitals) by the combined firm-university efforts, whereas the health biotech patents produced by firms separately or universities separately amount only to around 80% of their total biotechnology patent productions. The smallest share of health-related biotechnology patents (64%) is granted to the governmental institutions. These, on the other hand, account for a proportionally highest part (20%) of the agriculture and food related biotech patents. Universities are relatively less (8%) interested in doing environmental or industrial biotech research.

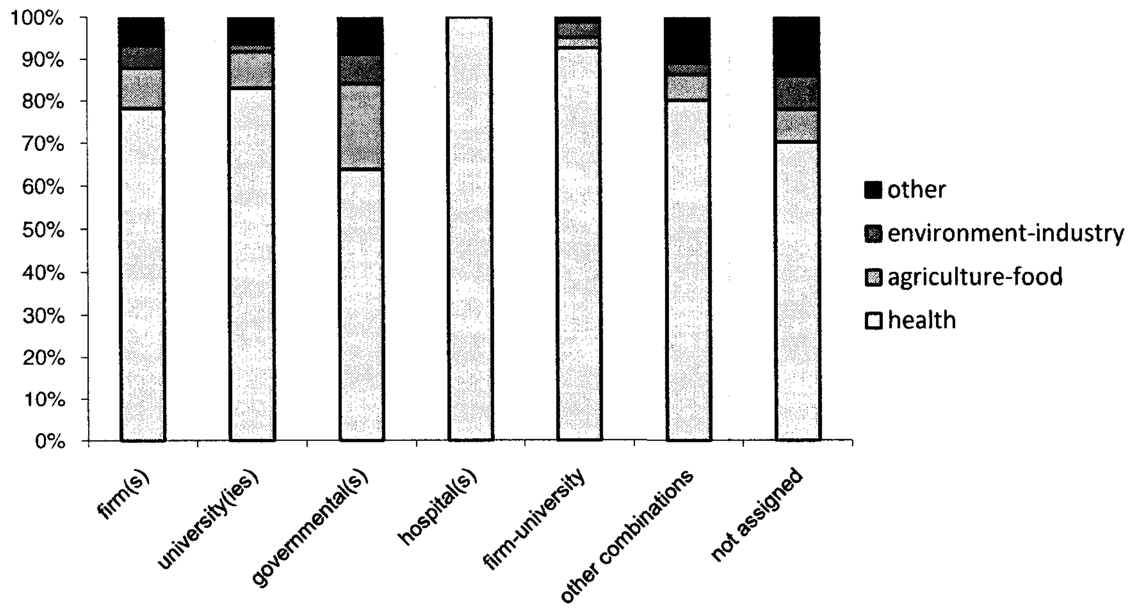


Figure 4-2: Proportions of patents by biotechnology field as granted to assignees in each category

Figure 4-3 shows the composition of the biotechnology fields in each cluster. Inside clusters, there seems to be an apparent focus on health-related biotechnology, whereas the patents produced outside the clusters are as often health as agriculture and food related. The highest focus on health-related biotechnology was found in the two most successful clusters, which largely disregard the agriculture and food related or environmental and industrial biotechnology. In general, the very low shares of patents in agricultural and environmental biotech fields are rather surprising, as it was expected to find evidence of some more specialized clusters (especially agriculture-related biotech in the Prairies – e.g. Saskatoon). Niosi (2003) however suggests that biotechnology firms in these fields stagnate or are in decline. The clusters with previously considerable shares of R&D in these fields have been observed to reorient themselves towards more profitable sphere of business in health-related products.

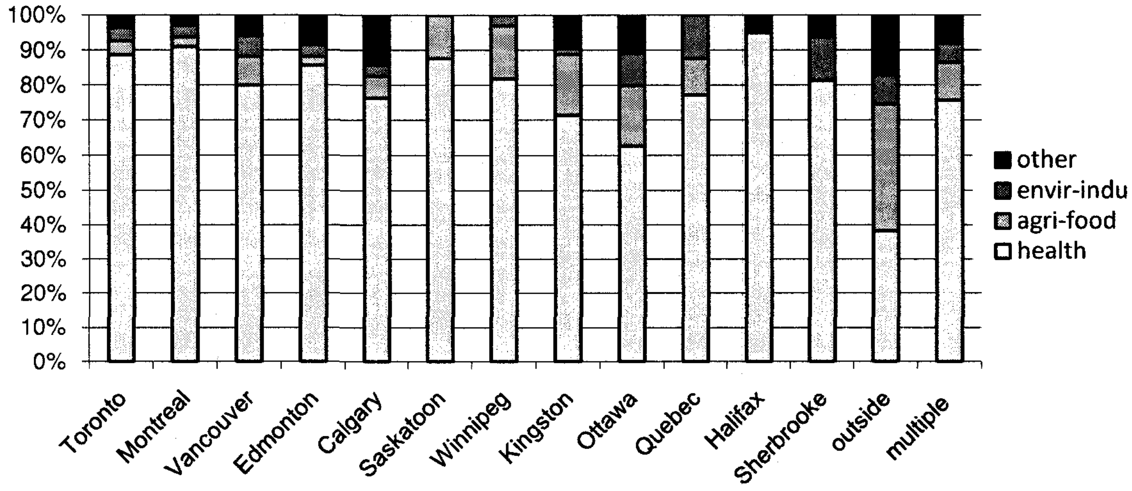


Figure 4-3: Biotechnology field composition of patents in each cluster

4.4 Collaboration in Canadian biotechnology based on patent co-assignment

Finally the collaboration propensity in Canadian biotechnology was examined. In order to trace the collaborative relationships among various entities the joint ownership of patents was explored, assuming that if a patent lists more than one assignee the invention has been developed under the active collaboration of the entities in question. Joint patent ownership is therefore considered to be a sign of the cooperation between institutions.[24] The analysis of assignments and co-assignments allowed understanding of the international, inter-cluster and intra-cluster collaborative patterns in biotechnology innovation.

Out of 3550 patents comprised in the database around 9% are not assigned and most of the patents (83%) have a single assignee, which does not show enough evidence of collaboration. The remaining patents (8%) are however assigned to several entities at the

---

[24] Joint ownership of patents was used previously to explore the inter-firm collaborations. For example, in order to investigate joint cooperative activities and formation of development coalitions, Gauvin (1995) used data on co-assignees of the patents granted by the Canadian government, Mariani (2000) examined co-patenting in the European chemical industry.

same time (multiple assignees). These patents were examined in more detail and the geographical aspect - the residences of the assignees and co-assignees - was specifically looked into. As Table 4-7 shows, most of the assignees in the database reside solely in Canada. Canadian entities are full or partial owners of around 70% of the USPTO biotechnology patents with at least one Canadian inventor. In 5% of patents Canadian assignees have foreign co-assignees. Most of these co-assignees (78%) reside in the USA, followed by France (4%) and Great Britain (3%). Also, 21% of patents in the database are fully assigned to a foreign entity, in most of these patents (77%) the foreign single assignee resides in the USA as well and is again followed by France (4%). Only very few patents are owned by the multiple assignees among which none resides in Canada.

Table 4-7: Patents by country of the assignees' residences

Assignees' residences	*number of patents*	*as % of all patents*	*claims (average)*
Only Canadian assignees	2310	65%	15.6
Foreign coassignees of Canadians	175	5%	16.0
Foreign single assignees	746	21%	17.7
Foreign multiple assignees	7	0.2%	16.0
No assignee	312	9%	14.4
All patents	**Σ 3550**	100%	**16.0**

With regards to the cooperation within Canada, it has already been shown in Table 4-1 that most of the biotechnology activity which takes place in Canada is concentrated inside clusters, usually the main ones. Only very few patents (1%) with co-assignees from multiple Canadian clusters, or from outside these clusters have been found. In addition to the very low level of inter-cluster patent ownership, only a marginal number of patents (2%) co-assigned to multiple entities within the clusters themselves was also observed.

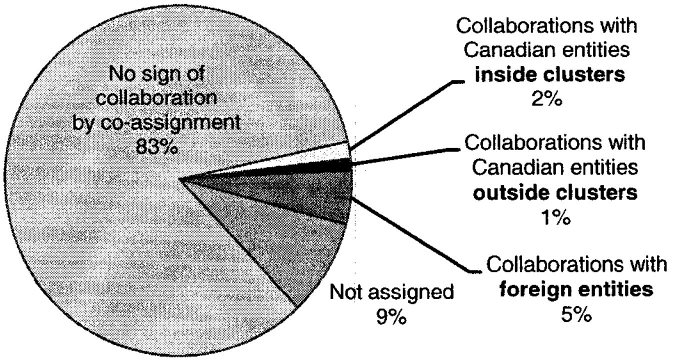


Figure 4-4: Collaboration pattern of Canadian biotechnology institutions as evidenced by the patent co-assignment

Based on all these findings the summarizing collaborative pattern of the institutions in Canadian biotechnology innovation was constructed. Figure 4-4 confirms that the amount of collaborative links with the US or other countries is surprisingly high in comparison with the apparently lacking joint biotechnology research in Canada. These findings however are not in agreement with the study of Gauvin (1995) who found that in Canada 78% of the joint patent ownerships (registered at CIPO) are domestic, while this figure would be only 34% for the used biotechnology sample. His database however included the patents across all the industries, and biotechnology may be a field with distinct collaborative patterns.

4.5 Conclusions

The results presented in this chapter confirmed that Canadian biotechnology is geographically highly concentrated. The majority of inventors reside in the three largest clusters of Toronto, Montreal and Vancouver. Several other agglomerations with sizeable patent production (Ottawa, Edmonton, Saskatoon, Calgary and Quebec) were identified together with some fairly small biotechnology concentrations (Winnipeg, Kingston, Halifax and Sherbrooke). The summary of the various characteristics of the eight most important Canadian biotechnology clusters is shown in Table 4-8.

Table 4-8: Summary of information on the seven most important clusters

		Toronto	*Montreal*	*Vancouver*	*Ottawa*	*Edmonton*	*Calgary*	*Sask.*	*Quebec*
# of patents		834	466	255	279	153	127	98	57
# of inventors		927	698	411	224	210	91	147	127
# of innovative firms		110	55	44	11	19	11	6	11
Patent quality				v. high			high	v. high	high
Firms' innovative productivity				low		low	v. high	low	low
Inventors' innovative productivity		high				high	v. high		
% of health-related biotech		89%	91%	80%	63%	86%	76%	88%	77%
Patent ownership structure	Firms	70%	56%	46%	19%	34%	52%	14%	30%
	Universities	10%	29%	47%	7%	30%	44%	66%	41%
	other (if share > 10%)	11% hosp.			62% gov.	21% gov.			11% gov.
# of prolific* firms		7	2		1				
# of highly prolific** universities			1	1		1	1	1	
# of prolific* hospitals		3							
# of prolific* gov. institutions		1			3	1			

*Prolific means here that the number of patents of the assignee is > 20 patents.

** Highly prolific means here that the number of patents of the assignee is > 50.

The findings of this chapter clearly suggest that biotechnology in Canada emanates from publicly-funded research. Universities are the most active institutions in Canadian biotechnology and the greatest producers of patents that are of high quality on average. They act as anchor tenants by attracting a pool of skilled workers and spin off new biotech firms. In small clusters in development, the local university is often nearly the only biotech patent producer in the cluster. In the larger and more mature clusters, where many firms are also located, the university's biotechnology activities represent a more modest share of the total biotech research. It was also noted that the production of patents is very different among Canadian universities. This is suggested to be related to two factors: First, it is the existence, quality and effectiveness of the technology transfer support present within these universities, consisting of the formal legal infrastructure and sufficient funds to file patents. Second, it also depends on the university IP rules and policies which stipulate whether the IP ownership is by default assigned to the university or may be retained by the individual inventors. As a consequence, several renowned research universities that are highly active in biotechnology research own only an inferior number of patents. The contribution of the government laboratories to the biotechnology research and development is also substantial. Around half of the Canadian biotechnology patents are owned by private companies. The patents assigned or co-assigned to firms are of higher quality than other patents on average.

This chapter has also examined the composition of biotechnology specializations in Canada. Biotechnology related to human health is the most significant biotechnology specialization in Canada in terms of number of firms, employment, R&D and revenues (Statistics Canada, 2007). It has been confirmed that health-related biotechnology represents by far the highest proportion of all biotechnology research innovation in Canada for all the various categories of assignees. In addition, the greatest and most successful clusters in Canada have a greater focus than the smaller ones on the health-related biotechnology field and largely disregard the agriculture and food related or environmental and industrial biotechnology. While the focus on the health-related biotechnology fields is overwhelmingly inside clusters, outside the clusters, however, the patents produced in Canada belong as often to the health related as to the agriculture and food related biotech specializations.

Based on the patent assignment and co-assignment data the intra-cluster, inter-cluster and international collaborative pattern in biotechnology innovation have also been constructed. Very little evidence of cooperation amongst Canadian biotechnology institutions, whether the collaborative ties lie within or outside clusters, was found. The most frequent typical partner for a Canadian biotechnology institution with which to pursue joint research activities is not another Canadian institution, but an institution abroad (mainly in the US). Further research is needed on this institutional cooperation.

Finally, it is not surprising that the inventions are often not owned by their creators. It was shown that the fruit of the inventive effort of the researchers is often claimed by universities, hospitals or companies. Moreover, although there is a great innovation capability among Canadian researchers, a lot of the intellectual property actually leaves the country. It has also an important implication for this research. Since the intellectual property policies of the various patent-owning institutions throughout the country are quite diverse, the information on the patent assignees often does not reveal the whole story behind the origin of the invention, its creation and the real innovative productivity of the location. Therefore in the following chapter it is intended to reach the roots of the inventive effort and focus on the real creators of innovations – the inventors themselves.

CHAPTER 5

GEOGRAPHICAL ASPECTS OF COLLABORATION

This chapter moves away from the previous focus on the patents and assignees, and concentrates rather on the inventors. Specifically, the collaboration patterns of the Canadian biotechnology inventors are analyzed and the geographical aspects of the collaboration examined. The chapter is based on the analysis of the collaboration instances[25] which are divided according to the location where they take place into the intra-cluster collaborations (both inventors in a collaborating pair are from the same cluster), inter-cluster collaborations (one of the inventors in a pair resides in a different cluster or elsewhere in Canada) and international collaborations (one of the inventors in a pair resides abroad). Figure 5-1 presents the overall collaboration pattern for the total of the Canadian biotechnology inventors. Well over half (60%) of the all collaboration instances take place inside the clusters and around 29% are distant ties directed abroad (mostly to the US). Only 11% of all the collaboration involves inventors from other Canadian clusters or from elsewhere in Canada. Most of the foreign collaborative ties are linked to the American inventors. These findings are slightly reminiscent of the results regarding the cooperation among institutions from the previous chapter (see section 4.4 of Chapter 4), where the cooperation with other institutions abroad was found to be much more common than the collaboration with other Canadian entities. Moreover, the most frequent foreign collaborating institution was American as well.

---

[25] An instance of collaboration (or simply collaboration) is a connection between a pair of inventors for the purpose of co-invention of one biotechnology patent.

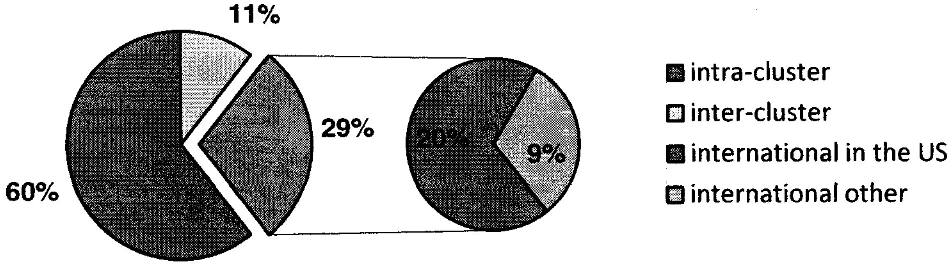
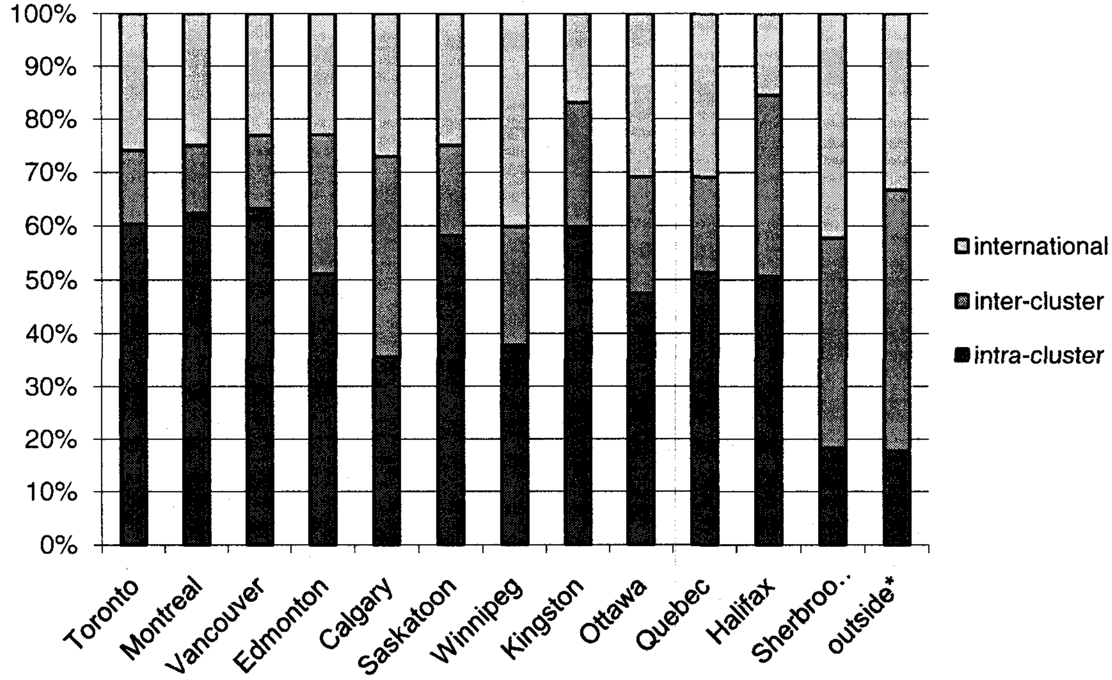


Figure 5-1: Collaboration pattern in Canadian biotechnology

In search for a more precise collaboration picture, exact proportions of the joint activities taking place within clusters (intra-cluster), among clusters (inter-cluster) and outside Canada (international) have been calculated for each cluster separately. Figure 5-2 shows that in general the inventors in three major Canadian clusters (Toronto, Montreal and Vancouver) have very similar collaborative patterns: more than 60% of collaborations between a pair of inventors take place within the cluster, where sufficient knowledge has probably been already accumulated, in around 25% of collaborations the expertise is sought abroad and only 13-14% of collaboration ties link the inventors with their partners in other clusters or elsewhere in Canada. As for the researchers outside these three major biotechnology agglomerations, the inventors in smaller clusters do not find all the needed expertise inside their own clusters and thus have to look for collaborators outside their cluster more frequently. The lowest share of collaborations inside the cluster is found for the inventors in the small cluster of Sherbrooke, but also for the cluster of Calgary. Figure 5-2 also confirms that if Canadian inventors decide to collaborate outside their clusters, they most commonly prefer to do so with inventors from abroad. In fact, Canadian researchers cooperate with their fellow inventors from other Canadian clusters much less frequently than expected. The preference of foreign over domestic collaborators is evident for the three main clusters which show the smallest percentage of collaborating pairs where each inventor comes from a distinct cluster. In some clusters however (Calgary, Edmonton, Kingston and Halifax), inventors who wish to collaborate outside their clusters still prefer to keep their collaborative ties inside Canada. While interpreting the figure, recall that it represents the proportions of the collaborations in each category and note that the total counts of instances of collaboration differ significantly among the clusters



*outside: inventors residing in Canada but outside the clusters

Figure 5-2: Collaboration pattern of the Canadian biotechnology clusters

In the remaining part of this chapter the results pertaining to each of the three collaborative locations will be presented in more detail. First, a bigger picture is shown by examining international collaborations in Canadian biotechnology, and then the results of the investigation of the inter-cluster collaborations inside Canada are presented.

5.1 International collaborations

In order to understand the geographical aspects of collaboration among inventors the vertices were grouped into several geographically-based classes. The vertices in the following two figures (Figure 5-3 and 5-4) represent all the inventors from the database grouped either by continents or by clusters. The link between each two groups represents the existence of a collaboration relation between them. The number above each link shows the total number of all instances of patent co-invention for all the members of each group, the strength of the lines represents the relative frequency of the cooperation.

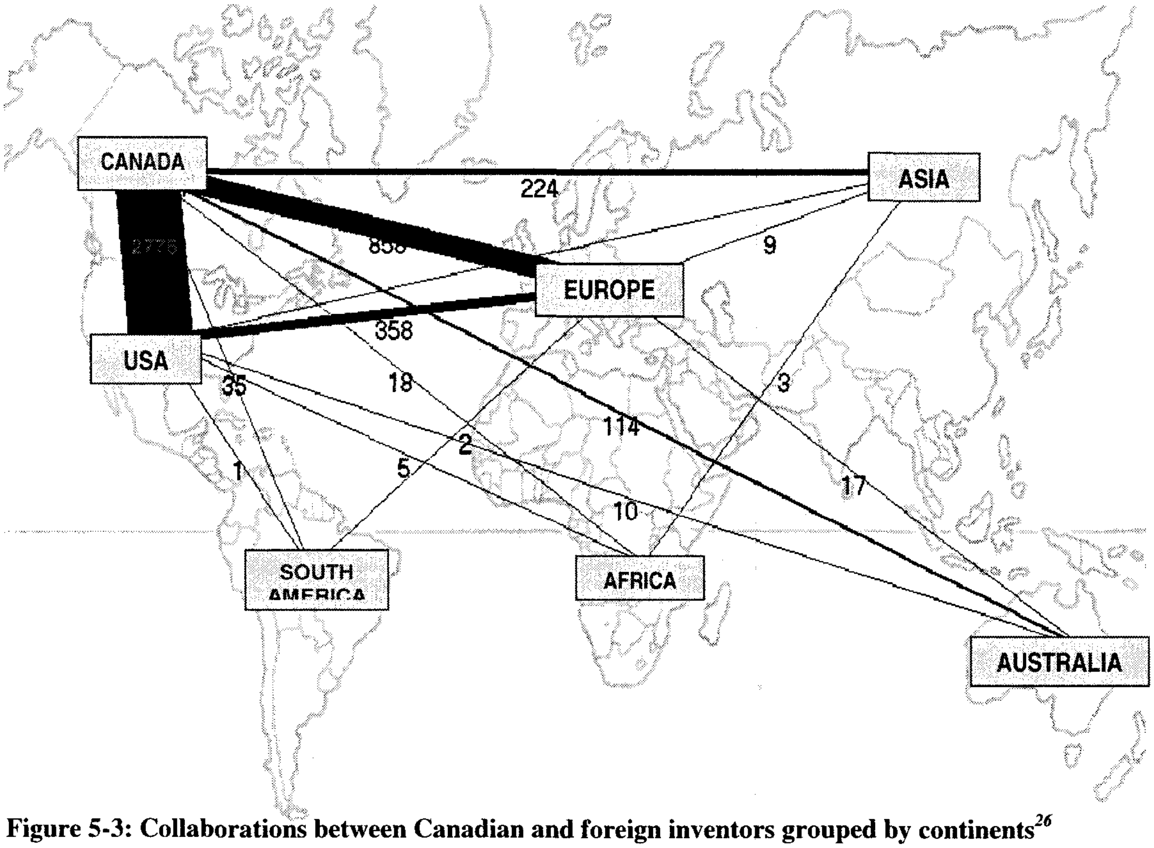


Figure 5-3: Collaborations between Canadian and foreign inventors grouped by continents[26]

---

[26] Recall that this is restricted database that does not account for all biotechnology patents in the world and consider the collaborations among the groups accordingly. Also, note Canada and the USA are separated into different groups in order to provide more information even though they evidently belong to the same continent.

As it was shown in Figure 5-1, 29% of all collaborative activities of Canadian biotechnology inventors are carried out across Canadian border. The collaborations between Canadian and foreign inventors grouped by continents are displayed in Figure 5-3. Around 21% of the international cooperation ties include European countries. Among them the most frequent collaborators of Canadian inventors are the French (8%) or the British (5%) inventors. Canada also works on the biotechnology patents with Australia (3%), Germany (2%) and Japan (2%). These results underestimate the collaboration intensity with inventors from European countries, since joint innovative activity between Canadian and European inventors would most probably be better shown by patents filed with the EPO or the CIPO.

Nonetheless, the majority (69%) of all foreign collaborations of Canadian inventors clearly takes place between Canada and the USA. Therefore more detailed geographical analysis of these partnerships has been carried out. Table 5-1 shows the absolute and relative numbers of collaborations among the biotechnology inventors residing in Canada and in the US regions. The most popular American cooperation partners for Canadian biotechnology inventors reside in the Northeast (32%) and Southwest (30%) regions. Among the US states, the highest number of Canadian cooperation links is directed towards California (27%) and the states in the North eastern region: New Jersey (10%), Massachusetts (9%), Philadelphia (6%) and New York (4%). The most popular collaboration partners in the Midwestern region reside in Michigan (4%) and Iowa (3%); in the South they come from North Carolina (5%) and Maryland (4%) and in the North western region they are mainly from Washington (3%).

Table 5-1: Number of collaborations among the inventors in Canadian biotechnology clusters and in the US regions (slightly modified US Census Regions)

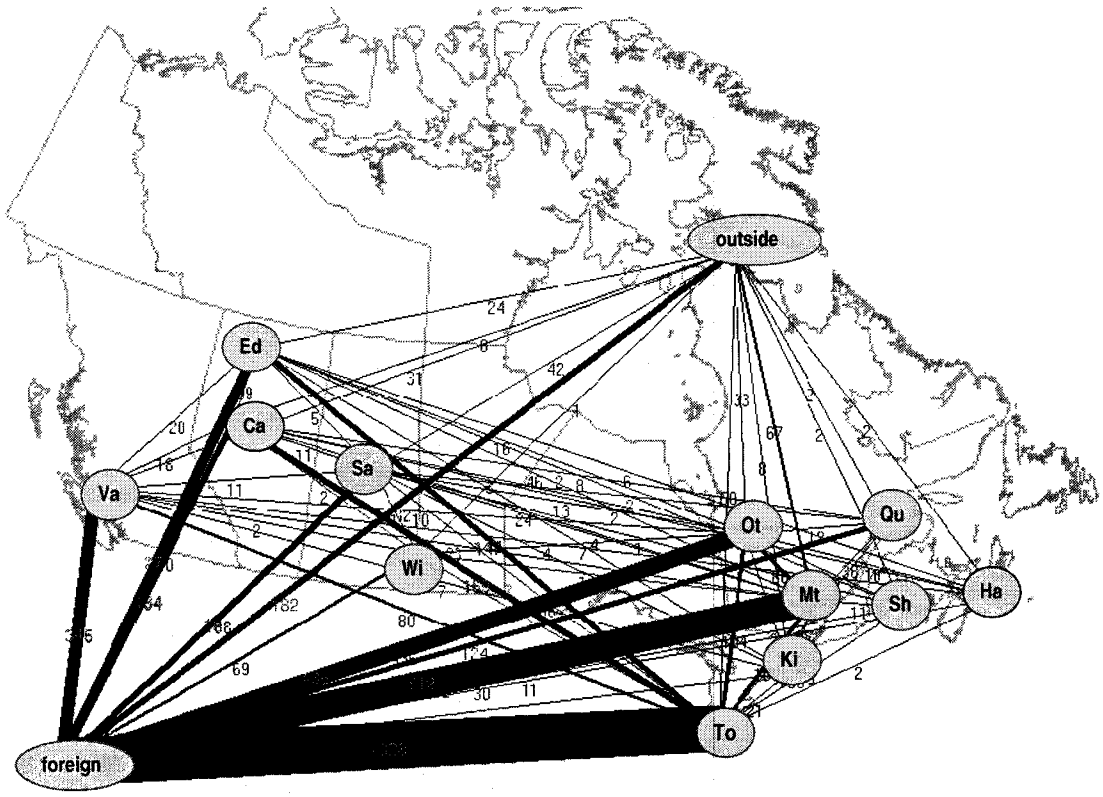
	Northwest	*Southwest*	*South*	*Midwest*	*Northeast*	***ALL USA***
Toronto	28 (3%)	**395 (38%)**	198 (19%)	218 (21%)	194 (19%)	***1033 (37%)***
Montreal	8 (1%)	130 (22%)	98 (17%)	74 (13%)	**272 (47%)**	***582 (21%)***
Vancouver	7 (2%)	**97 (35%)**	74 (27%)	36 (13%)	64 (23%)	***278 (10%)***
Edmonton	1 (1%)	**80 (63%)**	18 (14%)	9 (7%)	19 (15%)	***127 (5%)***
Calgary	5 (4%)	19 (15%)	17 (14%)	19 (15%)	**64 (52%)**	***124 (4%)***
Saskatoon		7 (23%)	**14 (47%)**	8 (27%)	1 (3%)	***30 (1%)***
Winnipeg		4 (12%)	6 (18%)	10 (29%)	**14 (41%)**	***34 (1%)***
Kingston		2 (10%)	5 (25%)	4 (20%)	**9 (45%)**	***20 (1%)***
Ottawa	29 (8%)	59 (17%)	50 (15%)	17 (5%)	**190 (55%)**	***345 (12%)***
Quebec		27 (29%)	14 (15%)	26 (28%)	26 (28%)	***93 (3%)***
Halifax		1 (14%)	**4 (57%)**		2 (29%)	***7 (0%)***
Sherbrooke		2 (29%)	**4 (57%)**		1 (14%)	***7 (0%)***
outside	1 (1%)	20 (21%)	18 (19%)	24 (25%)	**32 (34%)**	***95 (3%)***
ALL CANADA	*79 (3%)*	*843 (30%)*	*520 (19%)*	*445 (16%)*	*888 (32%)*	*2775 (100%)*

The table also shows the main collaboration partners per cluster. It is interesting to see that the Toronto inventors look for their collaboration partners most frequently in the geographically distant Southwest, while for the inventors from Montreal or Ottawa the most attractive collaboration deals are made in the close North eastern region. Even though the preferences of the western clusters of Vancouver and Edmonton for the western US states are not surprising, it is not at all obvious why the inventors in the western cluster of Calgary should choose to seal their partnership contracts predominantly in the eastern part of the US. These results suggest that once the deal cannot be made inside cluster or inside Canada the choice of the collaboration partner seems to depend much less on the geographical circumstances. But how important are the geographical selection criteria when searching for a collaborator inside Canada? The next section investigates the role of geography in the choice of a partner for joint research projects carried out within the Canadian borders.

5.2 Inter-cluster Collaborations

Figure 5-4 illustrates the collaborations among biotechnology inventors of different clusters inside Canada. To put the inter-cluster collaboration into perspective, international collaborations were included in the figure as well. The strength of the collaboration ties is shown both among individual Canadian clusters and between each cluster and all foreign countries grouped together. It could be easily observed that a great part of collaboration among biotechnology inventors takes place over the Canadian border. Canadian inventors rather pursue their joint research projects with inventors abroad, than with the ones from other Canadian clusters or outside these clusters, even if these reside relatively nearby. As was already discussed these foreign collaborating inventors are overwhelmingly from the US.

As it was shown in Figure 5-1, 11% of all collaborative activities take place among Canadian clusters. Figure 5-4 suggests that the strongest collaboration ties exist between the Toronto cluster and some other major clusters like Calgary, Edmonton, Montreal and Ottawa. Table 5-2 reveals a more detailed picture of the inter-cluster cooperation in Canadian biotechnology. Three collaborative patterns have been identified in Canadian biotechnology clusters. These were divided into: Eastern clusters with strong local partnerships, Eastern clusters with strong local and western partnerships and Western clusters with very strong Toronto partnerships.



Va	Vancouver	Ed	Edmonton	Ca	Calgary
Sa	Saskatchewan	Wi	Winnipeg	Ot	Ottawa
To	Toronto	Ki	Kingston	Mt	Montreal
Qu	Quebec	Sh	Sherbrooke	Ha	Halifax
outside	group of inventors residing in Canada, but outside the defined clusters				
foreign	group of non-Canadian inventors				

Figure 5-4: Collaborations among Canadian inventors grouped by clusters

Table 5-2: Number of collaborations among the biotechnology clusters in Canada

	Eastern clusters with strong local partnerships:			***Eastern clusters with strong local and western partnerships:***				***Western clusters with very strong Toronto partnerships:***				
	Montreal	*Ottawa*	*Quebec*	*Toronto*	*Kingston*	*Sherbrooke*	*Halifax*	*Vancouver*	*Edmonton*	*Calgary*	*Saskatoon*	*Winnipeg*
Montreal		**88 (29%)**	**38 (54%)**	**105 (15%)**	4 (6%)	**11 (39%)**		**23 (11%)**	**46 (15%)**	24 (9%)	4 (3%)	3 (8%)
Ottawa	**88 (21%)**		**12 (17%)**	**104 (15%)**	6 (10%)	**3 (11%)**	**10 (42%)**	10 (5%)	16 (5%)	3 (1%)	13 (10%)	**7 (18%)**
Quebec	38 (9%)	12 (4%)		15 (2%)						2 (1%)	2 (2%)	
Toronto	**105 (25%)**	**104 (34%)**	**15 (21%)**		**21 (34%)**	**3 (11%)**	2 (8%)	**80 (39%)**	**127 (42%)**	**145 (57%)**	**36 (28%)**	**22 (58%)**
Kingston	4 (1%)	6 (2%)		21 (3%)				7 (3%)	15 (5%)		1 (1%)	
Sherbrooke	11 (3%)	3 (1%)		3 (0%)					8 (3%)		1 (1%)	
Halifax		10 (3%)		2 (0%)				4 (2%)	6 (2%)	2 (1%)		
Vancouver	23 (6%)	10 (3%)		**80 (11%)**	**7 (11%)**		**4 (17%)**		20 (7%)	18 (7%)	11 (9%)	2 (5%)
Edmonton	46 (11%)	16 (5%)		**127 (18%)**	**15 (24%)**	**8 (29%)**	**6 (25%)**	20 (10%)		39 (15%)	5 (4%)	
Calgary	24 (6%)	3 (1%)	2 (3%)	**145 (20%)**			2 (8%)	18 (9%)	**39 (13%)**		11 (9%)	2 (5%)
Saskatoon	4 (1%)	13 (4%)	2 (3%)	36 (5%)	1 (2%)	1 (4%)		11 (5%)	5 (2%)	11 (4%)		1 (3%)
Winnipeg	3 (1%)	7 (2%)		22 (3%)				2 (1%)		2 (1%)	1 (1%)	
outside	67 (16%)	33 (11%)	2 (3%)	50 (7%)	8 (13%)	2 (7%)		31 (15%)	24 (3%)	8 (3%)	42 (33%)	1 (3%)
ALL	***413 (100%)***	***305 (100%)***	***71 (100%)***	***710 (100%)***	***62 (100%)***	***28 (100%)***	***24 (100%)***	***206 (100%)***	***306 (100%)***	***254 (100%)***	***127 (100%)***	***38 (100%)***

Eastern clusters with strong local partnerships: Three larger clusters in eastern Canada (Montreal, Ottawa and Quebec) pursue an expected collaborative behaviour, which is to look for the cooperation partnerships within a relatively short distance of their own cluster. Montreal's most frequent collaboration partners are from Toronto and Ottawa, as for Ottawa, these are from Toronto and Montreal and for Quebec it is mainly Montreal and Ottawa inventors. The inventors in these three clusters do not collaborate much with western Canada.

Eastern clusters with strong local and western partnerships: Toronto is an exception to the group of the larger eastern clusters, since the shares of its collaboration instances are quite evenly spread among all the most important clusters, whether they are geographically close as Montreal or Ottawa or they lie relatively far west as Calgary or Edmonton. The preferable direction of the Toronto inventors seems to be clearly towards the largest western clusters. The small eastern clusters of Kingston, Sherbrooke and Halifax usually find their collaboration partners in the relative geographical proximity (in Ottawa, Montreal and Toronto). In contrast to the larger eastern clusters, inventors in Kingston, Sherbrooke and Halifax do build their cooperation ties with the west much more often (in relative terms) than inventors from larger eastern clusters do.

Western clusters with very strong Toronto partnerships: The third collaboration pattern describes the typical cooperative behaviour of the western clusters of Vancouver, Edmonton, Calgary, Saskatoon and Winnipeg. For the inventors from these clusters most collaborative partners live in Toronto, while innovation partnerships from geographically closer clusters are usually much less attractive. Vancouver's biotech research partners come mainly from Toronto, but even Montreal is more preferred than closer clusters such as Calgary or Edmonton. The links with the highest number of collaboration instances in the whole inter-cluster collaborative network are the Toronto-Calgary link (145 collaborations) and Toronto-Edmonton link (127 collaborations). Calgary's and Edmonton's most important collaboration partner is by far the Toronto cluster, while the cooperation between each other is much more limited. The smaller western clusters of Saskatoon and Winnipeg follow a similar pattern - inventors in these clusters focus on collaboration with Toronto researchers. Both Saskatoon and Winnipeg also share one additional collaborative target, which the larger western clusters do not, and this is a well developed collaboration tie with Ottawa. Saskatoon cluster is also quite unique in that the highest share of its cooperative relationships is found outside the clusters.

Toronto is in total by far the most popular cooperation partner for Canadian biotechnology inventors from other clusters or elsewhere. 25% of all inter-cluster collaboration links in the whole network are directed towards the Toronto cluster. It is followed by Montreal (15% of links), Edmonton and Ottawa (both 11%). Vancouver seems to be less attractive partner for joint biotechnology research for Canadian inventors, since it accounts only for 7% of the collaborative links in the inter-cluster network. The conclusion stemming from this analysis is that the geographical distance is not likely to be the only critical factor when seeking partners outside the cluster. Other factors which are probably very decisive as well are the availability of particular inventors' biotechnology specialization and expertise, the size and reputation of biotechnology research, available facilities and funding, etc.

5.3 Distance-based analysis of all out-of-cluster collaborations

Given the specific geographical aspects of Canada (concentration of a great majority of its inhabitants along the southern border), the collaboration analysis based on political divisions (e.g., national versus international cooperation) does not actually tell a complete story about the distances between the collaboration partners. Many of the Canadian biotechnology clusters are located in a proximate distance from the US border and an international collaboration partner thus can be the closest one. For example, a Montreal inventor may find it much more convenient to establish collaborative partnership with his international counterpart in Boston than with a fellow Canadian inventor from Vancouver, since the distance is almost 10 times shorter. Therefore, all the out-of-cluster collaborations (including both international and inter-cluster ones) have been divided into four groups according to the distance between the residences of each collaborative pair: short range (distance < 600km), mid-range (600km < distance < 1600km), long range (distance > 1600km) and overseas (outside North America). Figure 5-5 shows the proportions of these collaborations for the inventors in each cluster. Out of the bigger clusters, Ottawa (58%) and Montreal (45%) have the highest percentages of short range collaborations, whereas the proximate cooperation projects do not seem to be popular in western clusters of Vancouver (5%), Saskatoon (6%), Edmonton (8%) or Calgary (16%). The low level of inter-cluster collaboration among the western clusters has already been suggested as well as their preference for the partners from Toronto and Southwest or Northeast US regions. The figure also shows the highest share of all clusters (58%-63) for their long-range partnerships. In most of the greater clusters, the proportions of the long-range and overseas collaborations are quite overwhelming, but the projects carried out over the mid-range distances do not seem to be that common.

All in all, almost 60% of all the out-of-cluster collaborations of Canadian inventors involve partners residing more than 1600km apart. Most of these distant partners live in Canada or the USA, but around one third of these collaborations link Canadians with overseas inventors. Mid-range collaborations are considerably less popular. Only around 13% of all collaborations outside cluster are carried out within 600km-1600km range. Much more frequent are joint research projects with geographically more proximate partners. In 28% of cases the out-of-cluster collaboration involves the partner located in the distance shorter than 600km.

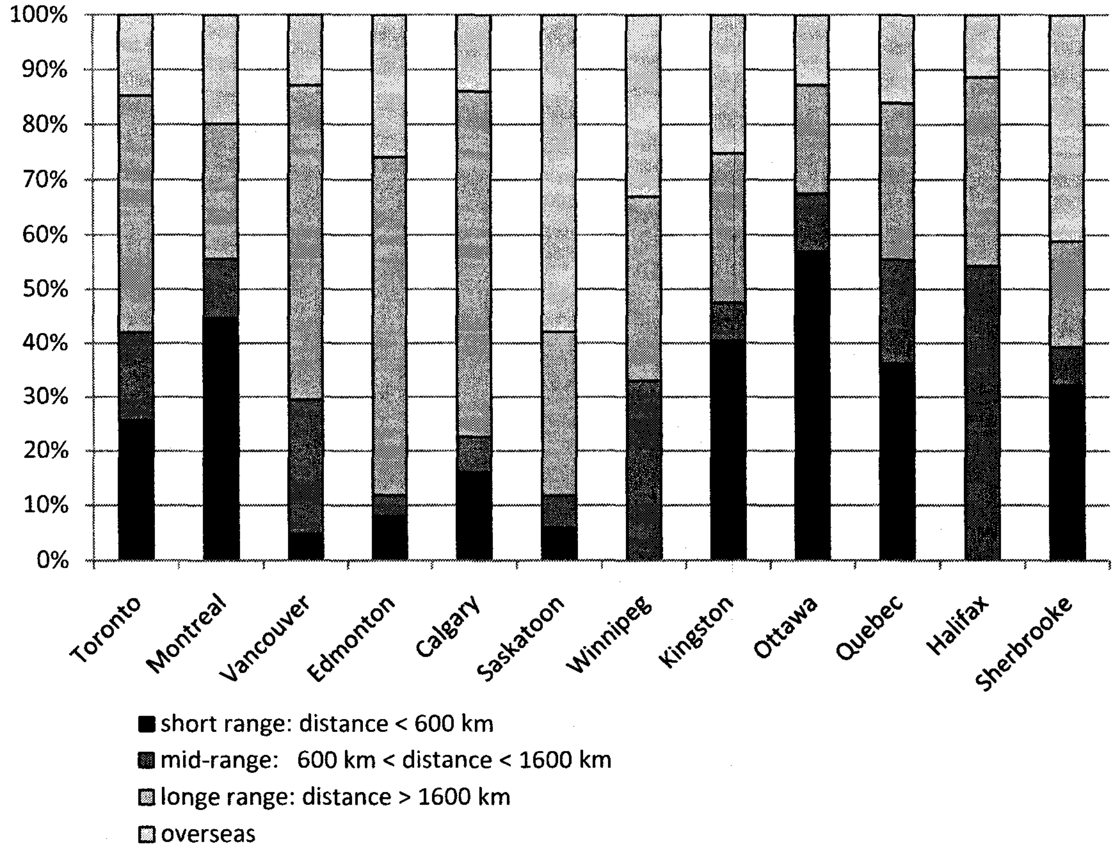


Figure 5-5: **Proportions of all out-of-cluster collaborations (including both international and inter-cluster cooperation) based on the distance between the collaborators**[27]

5.4 Conclusions

Geographical distance plays an important role when deciding on the partners for joint research projects in biotechnology. The results show that around 60% of the biotechnology collaborative activity which involves Canadian inventors takes place inside Canadian clusters.

---

[27] The distances are approximate: They are measured from the metropolitan centre of the Canadian clusters or from the geographical centre of the US states.

Canadian biotechnology inventors wishing to build cooperation ties outside their clusters were not found to collaborate very much with their fellow inventors from other Canadian clusters or elsewhere in Canada, even if these reside in a relatively close distance. The inter-cluster collaboration in Canada accounts on average only for 11% of all the collaborative ties. Three inter-cluster collaborative patterns have been identified in Canadian biotechnology innovation: Clusters with local partners, which are the bigger eastern clusters (Montreal, Ottawa, Quebec) that look for the cooperation ties within a relatively short distance of their own cluster. Clusters with both local and western partners are also situated in the eastern part of Canada (Toronto, Halifax, Sherbrooke and Kingston). They host inventors, who contract their partnerships for the most part with geographically closer inventors, but whose collaborative partners reside in the bigger western clusters as well. Clusters with Toronto partnerships are the western clusters (Vancouver, Edmonton, Calgary, Saskatoon and Winnipeg) and are characterized by their primary preference for the innovation partners from the relatively distant Toronto cluster. Toronto's inventors are in total by far the most popular cooperation partners for Canadian biotechnology researchers from other clusters or elsewhere in Canada.

Canadian inventors who decide to pursue their joint biotech research activities with inventors from outside their clusters most commonly prefer to search for their collaborative partners abroad. International ties account for the highest proportion of all the collaborations outside the clusters (29% of all cooperation links). The most popular foreign collaboration partners for Canadian biotechnology inventors reside south of the border, in the USA.

When the geopolitical divisions were disregarded and only geographical distances taken into consideration, it was observed that the distance plays an important role when deciding on the partners for joint research projects in biotechnology. An overwhelming preference of the Canadian inventors is towards local and relatively proximate partnerships. Nonetheless, if the suitable collaborators are not found within the distance of 600 km, the importance of the geographical factor significantly decreases, since in this case the inventors quite often opt for very distant or overseas cooperation. Other factors (biotechnology specialization, particular expertise, available facilities, previous acquaintance – e.g. former PhD supervisor, etc.) then become more prominent in explaining the inventors' choices.

Analogical analysis was carried out also for the collaboration among nanotechnology inventors (see Schiffauerova and Beaudry, 2008). The obtained results were very similar. An overwhelming preference of the Canadian inventors towards local and regional partnerships, especially within their own nanotechnology clusters, was also found to be present. Similarly, if the suitable collaborators could not be found within the region or at a short-range distance (600 km), the geographical criterion lost its importance. Inventors then quite often preferred very distant or overseas cooperation while disregarding the mid-range options.

CHAPTER 6

COMPARISON OF BIOTECHNOLOGY AND NANOTECHNOLOGY INNOVATION NETWORKS

6.1 Canadian biotechnology and nanotechnology clusters

In 2002, biotechnology was considered to be one the most dynamic and fast growing fields in Canada. According to the Statistics Canada (2005), biotechnology companies have more than quadrupled their revenues in 1997-2003. By 2002, Canada was the second most active country in the world in biotechnology in terms of new firms, venture capital and patents, after the US and ahead of the UK (Niosi, 2005). Metrics such as R&D spending, market capitalization as well as total number of firms and revenues all showed strong growth over the five years preceding 2002 (Ernst and Young 2002). Nevertheless, in the recent survey of Statistics Canada (2007) it was found that the number of innovative biotechnology firms increased only by 9% in the period of 2003-2005, whereas it increased by 31% between 2001 and 2003. Niosi (2006) noted that in recent years (particularly since 2000), Canadian biotechnology companies have experienced financing problems and even some of the well-financed firms have abandoned the field altogether. He suggests that the new trend of Canadian biotechnology is directed towards concentration of activity into a small number of dedicated biotechnology companies. Figure 6-1 shows the growth of biotechnology and nanotechnology patents in Canada based on the year of granting. It illustrates the phenomenal growth by the steeply increasing annual numbers of patents in those years. It is also evident that after the peak in 2001-2002 the number of biotechnology patents invented or co-invented by Canadians has been decreasing.

The research on nanotechnology was rather sporadic until 1987 when the annual acceleration of the patent production rate started. Apart from a short period of decline in

2000, the number of nanotechnology patents granted per year has been steadily increasing and during the last 15 years it has in fact increased ten-fold. Moreover, in 2004 the annual growth of the granted nanotechnology patents is almost reaching the biotechnology patents annual growth.

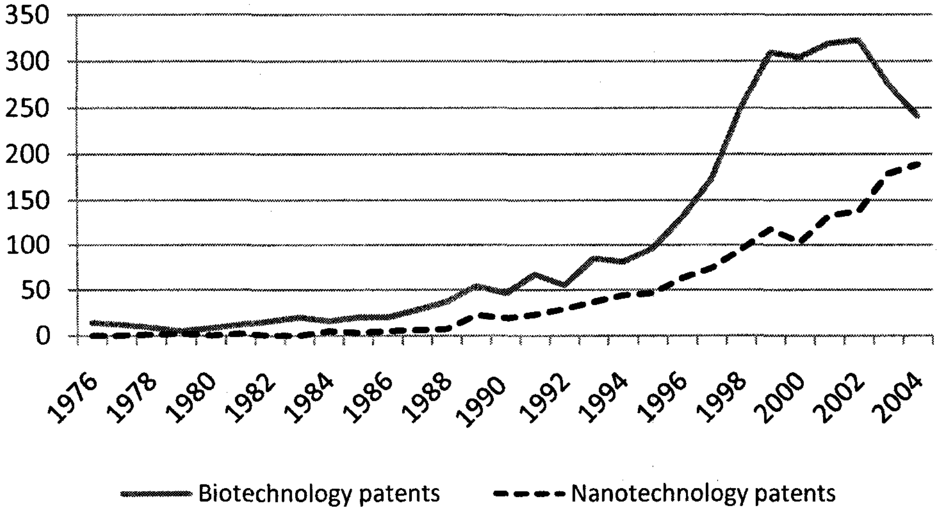


Figure 6-1: Patents of Canadian biotechnology nanotechnology inventors by the year of granting

The production of both biotechnology and nanotechnology patents is however not uniform throughout Canada. Most of the Canadian biotechnology or nanotechnology innovation is concentrated in only several regions. Based on the residences of inventors 12 Canadian biotechnology clusters and 8 Canadian nanotechnology clusters have been identified. As described in Chapter 4, 20% of biotechnology inventors reside in the Toronto cluster (34% of all patents), 15% in the Montreal cluster (19% of all patents) and 9% in the Vancouver cluster (10% of all patents). Only a small portion of biotech inventors (4%) residing in Canada lives outside the defined clusters (2% of patents) and around 29% of the innovators in this sample reside outside the Canadian borders (21% of all patents are assigned solely to foreigners).

The situation is quite different in nanotechnology. The greatest part of all the patents (47%) invented or co-invented by Canadian scientists is assigned to the foreign entities, most of which reside in the US; 69% of the patents owned by non-Canadian subjects is assigned to a single American company – Xerox Corporation. Only 28% of the inventors whose patents were assigned to foreign subjects are foreigners as well, most of them (62%) reside in the Toronto cluster. The consequence is a low number of assignees compared to a disproportionally high number of inventors residing in Toronto (see Figure 6-3). As for the number of inventors residing in each of 8 identified Canadian nanotechnology clusters, Toronto cluster is leading (25% of inventors), followed by Montreal and Ottawa (9% of inventors in each cluster).

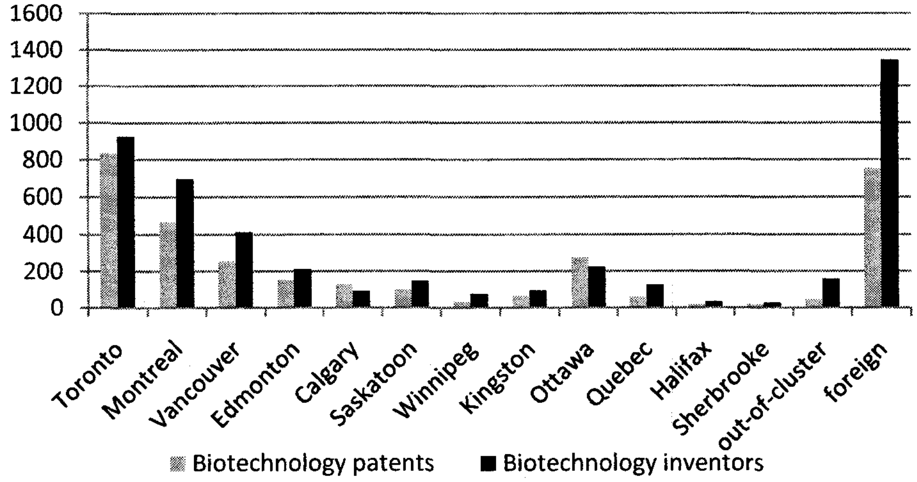


Figure 6-2*:* **Patents and inventors in each biotechnology cluster based on the location of patent`s assignees and the residences of inventors**

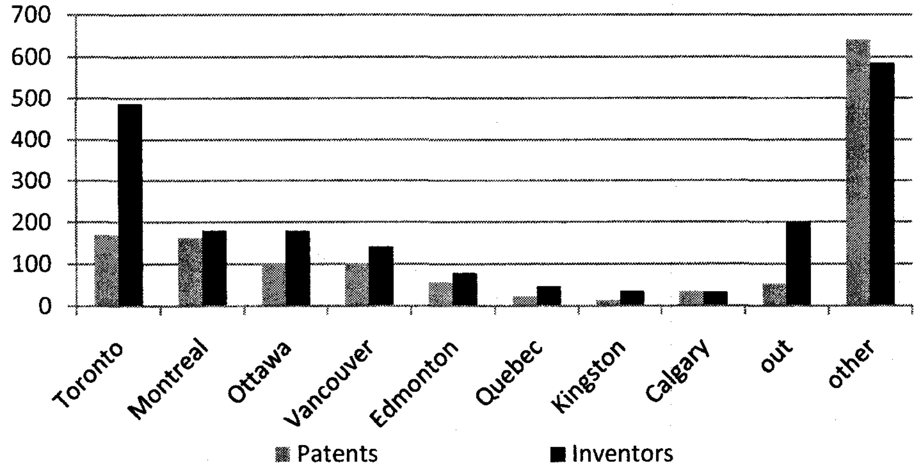


Figure 6-3*:* **Patents and inventors in each nanotechnology cluster based on the location of patent`s assignees and the residences of inventors**

Figure 6-2 and Figure 6-3 show the respective situations of the 12 biotechnology and 8 nanotechnology clusters, as described by the measures of the number of the patents in the cluster and the number of inventors.

Table 6-1 confirms that most of the Canadian biotechnology and nanotechnology activities take place within clusters, usually the few main ones. Only around 1% (biotechnology) or 4% (nanotechnology) of the patents are owned by assignees with residences in Canada but outside the predefined clusters. In both technologies there are only very few patents with co-assignees from multiple Canadian clusters. As was already discussed in Chapter 4 (see section 4.4), the lack of common inter-cluster ownership of patents suggests that there is not much cooperation at the assignee level between clusters, and if there is, ownership of patents is not shared.

The third column shows the numbers of patents per inventor produced in various biotechnology clusters (counted as the number of patents co-invented by at least one inventor from the cluster divided by the number of inventors who at least once patented while living in that cluster) and gives a certain indication about the productivity of the inventors in each cluster. The highest biotechnology productivity is in the Calgary cluster (2.19 patents per inventor), followed by Toronto (1.44 patents per inventors). As for the nanotechnology clusters (see the sixth column), this number is again highest for the Toronto cluster (1.52 patents per inventor). In the seventh column an alternative indicator based on the nanotechnology assignee's residence (counted as the number of patents allocated to the clusters by assignees' residences divided by the number of inventors allocated to that cluster based on their most frequent residence) has been computed. It is extremely low for the Toronto cluster (0.35 patents per inventor), to which only very little patents are assigned, even though it has many inventors. This suggests that many nanotechnology inventors residing in Toronto work for companies headquartered in the US. As has already been mentioned above, 62% of inventors whose patents were assigned to the foreign subjects reside in the Toronto cluster.

Table 6-1: Summary of the results for biotechnology and nanotechnology clusters

High technology cluster	***Biotechnology***			***Nanotechnology***			
	Number of patents[a]	*Number of inventors*[c]	*Patents per inventor*[d]	*Number of patents*[b]	*Number of inventors*[c]	*Patents per inventor*[d]	*Patents per inventor*[e]
Toronto	842	927	1.44	169	487	1.52	0.35
Montreal	469	698	1.05	162	180	1.16	0.90
Vancouver	258	411	0.95	103	142	0.95	0.73
Ottawa	286	224	1.26	103	179	0.92	0.58
Edmonton	158	210	1.21	57	79	0.94	0.72
Calgary	129	91	2.19	34	33	1.30	1.03
Quebec	61	127	0.97	23	47	0.79	0.49
Kingston	64	94	1.01	14	35	1.05	0.40
Saskatoon	101	147	1.04	-	-	-	-
Winnipeg	34	77	0.91	-	-	-	-
Halifax	20	33	1.06	-	-	-	-
Sherbrooke	16	26	1.07	-	-	-	-
out-of -cluster	47	159	1.16	52	201	0.87	0.26
foreign residence	753	1345	-	640	585	-	-
non-assigned[f]	312	-	-	86	-	-	-
Canada	***Σ 3550***	***Σ 4569***	*1.21*	**Σ 1443**	**Σ 1968**	*1.03*	*0.69*

[a] Based on the residence of the assignees. Notice, that the numbers of patents per cluster are different from Table 4-1 in Chapter 4, where the main focus was on the assignees. The patents with multiple assignees belonged to the special category, which was further discussed. Here, the number of patents per cluster is compared and the geographical aspects are the main concern. Therefore, the patents with multiple assignees from multiple clusters (37 patents) were allocated to only one cluster.

[b] Based on the residence of the assignees (the patents with multiple residences were allocated to only one cluster)

[c] Based on the residence of the inventors (the inventors who patented while living in several clusters were assigned to only one cluster)

[d] Counted as the number of patents co-invented by at least one inventor from the cluster divided by the number of inventors who at least once patented while living in that cluster.

[e] Counted as the number of patents allocated to the clusters by assignees' residences divided by the number of inventors allocated to that cluster based on their most frequent residence

[f] The inventor still has not decided who will own the patenting rights

6.2 Collaboration patterns in Canadian biotechnology and nanotechnology

The following two sections explore the collaboration characteristics and the structure of innovation networks formed by the inventors in the clusters. The network of

Canadian biotechnology inventors which was created includes 4569 vertices (representing inventors) and 9731 edges (representing collaborative relations[28]), whereas the network of Canadian nanotechnology inventors involves only 1968 vertices and 4920 edges. The main concern consists in the study of knowledge flows and information exchange among the researchers, i.e. in the characterization of the links between them. For instance, it was found that 36% (biotechnology) or 34% (nanotechnology) of all collaborative relations between pairs of inventors involve repetitive instances of collaboration[29]. In some cases the cooperative relationships actually seem to be very fruitful, as the most frequent collaboration between a pair of inventors was repeated 60 times (biotechnology) or 50 times (nanotechnology). Most of the relationships between a pair of inventors are, however, single collaboration instances (i.e., they resulted in only 1 patent).

An inventor in Canadian biotechnology network has on average 4.26 collaboration partners[30] (5 partners in the nanotechnology network), but some of them have a considerably higher number of relationship ties, the highest one amounting to 66 co-inventors (54 co-inventors for nanotechnology). Canadian inventors most commonly have one collaborator (16% of biotech inventors and 12% of nanotech inventors), two collaborators (20% of biotech and 19% of nanotech) or three collaborators (17% of biotech and 16% of nanotech). Only a small amount of inventors (4% in both networks) do not collaborate with anybody else on their patent(s) (single inventors or isolates), and only a few (6% of biotech and 8% of nanotech inventors) have more than 10 co-inventors. The average number of collaborating partners per inventor and per patent in

---

[28] Each collaborative relation (also called a tie or a link) represents a connection between a pair of inventors, which involves one or more instances of co-invention of a biotechnology patent.

[29] Recall, that an instance of collaboration (or simply collaboration) is a connection between a pair of inventors for the purpose of co-invention of one biotechnology patent. Each collaborative relation may thus involve one or more instances of collaboration (collaborations).

[30] Collaboration partner (or collaborator) is here defined as a co-inventor of at least one biotechnology patent registered at the USPTO.

each cluster is presented in Table 6-2. While in biotechnology these numbers seem quite comparable for each cluster, in nanotechnology the average number of collaborators in the Toronto cluster clearly stands out (it is almost double compared to other clusters). This suggests that the Toronto nanotechnology inventors collaborate more intensively and exchange information with more inventors than researchers in other clusters.

The general results of this thesis (4.26 or 5 collaborators per inventor) are comparable with the average number of collaborators per inventor found by Beaucage and Beaudry (2006) who observed 5.12 collaboration partners per Canadian biotechnology inventor. Even though their figures are slightly higher, they roughly correspond to ours in terms of the average collaboration partners in each biotechnology cluster. Out of the three main biotechnology clusters which they studied, the average Montreal inventor has the highest number of collaborators while the Toronto inventor the lowest (which can be observed in the results obtained for biotechnology in Table 6-2 as well). The average number of collaborators per inventor for the networks of Balconi *et al.* (2004, calculated from p.139, Table 5) was calculated in order to compare its value with the obtained results. The calculation shows that the networks of Balconi *et al.* (2004) have on average 2.09 collaborators per inventor, considerably less than the 4.26 collaborators (biotechnology) or 5 collaborators (nanotechnology) observed in the networks of this thesis. The difference can be explained by the distinct samples of patents selected for the analysis: Contrarily to the narrowly focused patent sample used in this thesis (only biotechnology or nanotechnology), in the study of Balconi *et al.*, the industry range is quite broad. Newman's findings (2001a) differ even more from these results. He observed a much larger number of collaborators in his innovation networks; especially for the scientists in experimental disciplines (for instance, an average high-energy physics scientist had 173 collaborators during a five year period). The scientific papers have however traditionally more numerous co-authors than the patents (the largest number of co-authors on a single paper found by Newman was 1681!), since joint article authorship was found to reflect a variety of phenomena other than the exchange of information and research collaboration.[31] Even though the legal requirements for article co-authorship and patent co-inventorship are officially very similar, the number of article co-authors is on average much higher than the number of co-inventors of the patent which reflects exactly the same discovery or invention. Ducor (2000) found that the number of article co-authors is on average more than three times higher than the number of inventors on the corresponding patent.

Table 6-2 also shows the results of some basic statistics regarding collaborators and collaborations in clusters. The results in the second and the fifth columns (co-inventors per patent) would at first glance suggest that the average team size is similar in all the clusters; however ANOVA tests (see Appendix E for biotechnology and Appendix F for nanotechnology) showed that the population means are in fact different and the team sizes within both the Canadian biotechnology and nanotechnology research thus differ across the country. Balconi *et al.* (2004) proposed that the differences in team sizes may be explained by the affiliations of the inventors – the researchers affiliated to the academic institutions work in larger teams and for a larger number of applicants than do industrial researchers. This research does not yet distinguish between academic and industrial researchers and to validate this hypothesis for Canadian biotechnology and nanotechnology, but there is an intention to do so in future. The third and the sixth columns in Table 6-2 show a number of collaborative instances per inventor in each cluster. To sum up, Table 6-2 suggests that in order to generate innovations, biotechnology researchers in the clusters of Saskatoon, Ottawa, Edmonton and Montreal collaborate slightly more intensively and exchange information with more inventors than researchers in other clusters. In nanotechnology, it is mainly the inventors from the Toronto cluster that show substantially higher collaborative intensity.

---

[31] Cockburn and Henderson (1998) suggest that article co-authorship may be offered as a quid pro quo for supplying information or resources, it can serve as a means of resolving disputes about priority, it may also be an acknowledgement of an intellectual debt, it may just be listing of laboratory directors or other project leaders as authors or it may reflect an effort to gain legitimacy, or admission to networks of other researchers.

Table 6-2: Statistics regarding collaborators or collaborations for each cluster

	Biotechnology			***Nanotechnology***		
High technology cluster	*Collaborator per inventor*	*Co-inventors in one patent*	*Collaboration per inventor*	*Collaborators per inventor*	*Co-inventors in one patent*	*Collaborations per inventor*
Toronto	3.49	2.91	8.88	7.02	2.88	14.13
Montreal	4.04	3.23	7.57	3.97	3.07	7.48
Vancouver	3.67	3.05	5.95	3.65	2.89	5.22
Ottawa	4.55	2.96	9.28	3.69	2.88	5.34
Edmonton	4.83	3.28	8.49	4.10	3.23	6.35
Calgary	3.93	2.59	10.13	3.88	3.41	8.55
Quebec	3.31	2.83	4.78	3.49	3.78	4.81
Kingston	2.86	2.68	4.52	3.71	3.43	6.63
Saskatoon	4.54	3.29	8.10	-	-	-
Winnipeg	2.53	2.11	3.08	-	-	-
Halifax	2.09	2.25	3.24	-	-	-
Sherbrooke	2.50	2.44	3.23	-	-	-
out-of cluster	2.85	2.96	4.05	2.93	2.62	3.61
Average in Canada	***4.26***	***3.09***	***7.46***	***5.00***	***3.00***	***8.60***

6.3 Local collaboration in the cluster-based subnetworks

It has been suggested and empirically supported that firms in clusters are more innovative (Baptista and Swann, 1998; Beaudry, 2001; Beaudry and Breschi, 2003; Beaudry and Swann, forthcoming). The companies collocated in a close geographical proximity enjoy numerous benefits discussed in the literature review. Biotechnology and nanotechnology knowledge is largely tacit, which limits knowledge diffusion over long distances. In fact, the transmission of tacit information and knowledge spillovers is usually associated with face-to-face contact. Collaboration among the inventors working in biotechnology and nanotechnology clusters is thus strongly encouraged by the benefits of acquiring knowledge which the subjects located within short geographical distance spill over.

This section of the thesis analyzes these local collaborations carried out entirely within clusters. Both Canadian biotechnology and nanotechnology innovation networks have been divided into geographically based subnetworks, where each subnetwork strictly includes inventors who reside in one particular cluster, while excluding the ones that do not. Out-of-cluster and foreign inventors are therefore eliminated for the time being. For each of the subnetworks created in this manner several network characteristics were calculated. The remaining sections of this chapter briefly discuss several of the basic structural properties of the network and explain the indicators used in this thesis to measure them. It is shown how these characteristics could be related to efficiency in the knowledge diffusion among the inventors within the clusters and the possible impact on innovation creation in the cluster is suggested.

6.3.1 Collaboration characteristics in the subnetworks

As Table 6-3 shows, 18-50% of collaborative relations between pairs of biotechnology inventors residing in the same cluster (and 20-47% between nanotechnology inventors) involve repetitive instances of collaboration. Biotechnology inventors in Toronto and Calgary tend to pursue collaborative relations with the same partners more often than the biotechnology inventors in Montreal, Vancouver or Edmonton. In Halifax, half of the collaborative ties of the local biotechnology inventors include repetitive collaborative relationships. As for the nanotechnology inventors, those in Toronto, Montreal and Ottawa collaborate with the same partners much more often than inventors in Vancouver or Edmonton. With regards to the smaller nanotechnology clusters, in Kingston almost half of the collaborative ties of the local inventors include repetitive collaborative relationships and the repetitiveness is also high in Calgary.

Table 6-3: Collaboration characteristics in biotechnology and nanotechnology cluster-based subnetworks

	Biotechnology			***Nanotechnology***		
Cluster-based subnetwork	*Number of collaborating pairs*	*% of repeate collaboration*	*Max number of repeated collaborations*	*Number of collaborating pairs*	*% of repeated collaborations*	*Max number of repeated collaborations*
Toronto	1120	43%	60	1295	38%	50
Montreal	1027	36%	11	201	36%	29
Vancouver	568	37%	10	199	20%	12
Ottawa	343	36%	19	218	35%	6
Edmonton	334	37%	14	112	24%	6
Calgary	91	41%	16	41	41%	8
Quebec	155	18%	7	53	21%	3
Kingston	96	33%	10	36	47%	4
Saskatoon	259	28%	8	-	-	-
Winnipeg	54	19%	3	-	-	-
Halifax	20	50%	5	-	-	-
Sherbrooke	10	20%	3	-	-	-
Network	***9731***	***36%***	***60***	***4920***	***34%***	***50***

In biotechnology and nanotechnology networks, the strongest collaboration link in the network, i.e. the most frequently repeated collaborative relation, concerns two inventors in Toronto (one in biotech and one in nanotech). They repeated their collaboration 60 (biotech) or 50 (nanotech) times. In smaller clusters, the maximum number of repeated collaborations is lower. Within the biotech subnetworks it is still relatively low for the larger clusters of Montreal (11) and Vancouver (10), where on average, innovative activities involve slightly more co-inventors who collaborate with each other less often. The maximum number of repeated collaborations in the nanotechnology subnetworks is relatively high for Montreal (29), but surprisingly low for the similarly-sized Ottawa (6) and somewhat smaller Vancouver (12). On average, the innovative activities of nanotechnology inventors in Toronto involve considerably more co-inventors who collaborate with each other more often than in any other nanotechnology cluster studied.

6.3.2 Fragmentation of the subnetworks

In order to assess the fragmentation of the subnetworks the network components were identified and their major characteristics determined (see Table 6-4). A component is defined as the maximal connected subnetwork (Wasserman and Faust, 1994). It is a part of the network which includes a maximum number of vertices which are all directly or indirectly connected by links. The largest component (in absolute value) of the biotechnology subnetworks is found in the Montreal cluster (109 interconnected inventors, which comprises 16% of inventors), even though Toronto has almost twice as many inventors (the largest component size is 98, which is only 11% of inventors). In nanotechnology, the Toronto largest component consists of 155 inventors, or around 32% of all the Toronto nanotechnology inventors. Even though the Toronto nanotechnology cluster thus shows a surprisingly high interconnectedness of the inventors, the rest of the clusters have relatively much smaller largest components.

The second largest components of the biotechnology subnetworks in Montreal and Toronto are of similar sizes, with that of Vancouver being much smaller. The cluster of Vancouver is in general more fragmented than the other two. In Saskatoon, even the second largest component is composed of proportionately many inventors. Regarding the nanotechnology subnetworks, the second component in Toronto is more than 10 times smaller than the first largest component. In contrast, for the other three nanotechnology clusters the second largest components are around half the size of the largest ones and in Edmonton, they are almost of the same size. These nanotechnology cluster subnetworks are overall more fragmented than the Toronto one.

The average component size is fairly small for all the biotechnology and nanotechnology clusters (around 2-3 inventors). As hinted by the previous paragraph, Saskatoon, which comprises components of a large relative size, scores the highest on the average number of interconnected inventors (4.32 inventors). The second rank is occupied by Montreal and Ottawa (both have on average 3.2 connected inventors), but

Toronto has a mean of only 2.71 inventors in a component. In nanotechnology, on the other hand, the Toronto subnetwork stands out: it includes components of a larger relative size: the mean number of interconnected inventors is 3.69. Moreover, half of the inventors form only around 20% of all the components, whereas this percentage is much larger for all the other nanotech clusters. The remaining nanotechnology subnetworks have on average a comparable numbers of connected inventors (around 2.5-2.9 inventors).

The counts of isolate vertices in both biotechnology and nanotechnology subnetworks are proportionately comparable for the large clusters (15%-19% of all the vertices) and relatively high for the smaller clusters (e.g., in Sherbrooke almost half of the biotech inventors are isolated). The Toronto nanotechnology subnetwork has the lowest percentage of isolate vertices (11%) of all the clusters. Many inventors have collaborators outside their clusters or outside Canada that contribute to linking indirectly inventors from the same cluster. As explained previously, only cooperation based on close personal contacts, which are limited by geographical distance, is considered here.

Taken all of the above in account, it can be concluded that the biotechnology network seems to be slightly less fragmented than the nanotechnology one. However, in nanotechnology there appears to be a well interconnected network component in Toronto, but the rest of the Canadian nanotechnology inventors are working in a relatively disconnected groups. Even when the full network values are considered while disregarding the geography, the average component size in nanotechnology is somewhat smaller (5.11 in biotech and 4.84 in nanotech), while the share of the components which include 50% of all the inventors is much higher in nanotechnology network (only 10% in biotech but 26% in nanotech networks). The percentage of isolates is comparable in both networks (4% for both biotechnology and nanotechnology).

This result was expected. The specialization fields within the biotechnology are quite close in their scientific nature and are often overlapping. The inventors in the biotechnology network should thus be more interconnected between each other. Nanotechnology, on the other hand, includes many quite disparate fields, where the inventors understandably work in more separated groups. Nanotechnology would therefore appear more as a brand name than a “single” technology so far.

Table 6-4: Fragmentation in biotechnology and nanotechnology cluster-based subnetworks

	Biotechnology							*Nanotechnology*						
Cluster-based subnetwork	*Number of components*	*1st largest component*		*Ratio 2nd/1st*	*Average comp. size*	*Share of comp. with 50% inventors*	*Isolates as % of inventors*	*Number of components*	*1st largest component*		*Ratio 2nd/1st*	*Average comp. size*	*Share of comp. with 50% inventors*	*Isolates as % of inventors*
		Size	*As %*						*Size*	*As %*				
Toronto	342	*98*	*11%*	0.36	2.71	13%	19%	132	155	32%	0.08	3.69	21%	11%
Montreal	218	109	16%	0.31	3.20	11%	15%	73	21	12%	0.48	2.47	34%	19%
Vancouver	134	38	9%	0.47	3.07	15%	16%	52	21	15%	0.43	2.73	31%	18%
Ottawa	70	75	33%	0.15	3.2	9%	17%	68	23	13%	0.43	2.63	29%	18%
Edmonton	67	49	23%	0.43	3.13	11%	17%	27	15	19%	0.93	2.93	26%	15%
Calgary	39	15	16%	0.60	2.33	18%	24%	17	8	24%	0.50	1.94	35%	30%
Quebec	44	11	9%	0.82	2.89	21%	15%	19	6	13%	1.00	2.47	47%	17%
Kingston	38	8	9%	0.75	2.47	24%	18%	17	5	14%	1.00	2.06	53%	29%
Saskatoon	34	54	37%	0.67	4.32	6%	13%	-	-	-	-	-	-	-
Winnipeg	44	7	9%	0.86	1.75	25%	36%	-	-	-	-	-	-	-
Halifax	20	6	18%	0.67	1.65	30%	39%	-	-	-	-	-	-	-
Sherbrooke	18	3	12%	1.00	1.44	33%	46%	-	-	-	-	-	-	-
Network	***894***	***579***	***13%***	***0.32***	***5.11***	***10%***	***4%***	***407***	***336***	***17%***	***0.09***	***4.84***	***26%***	***4%***

6.3.3 Structural cohesion of the subnetworks

Structural cohesion refers to the degree to which vertices are connected among themselves. The most common measure of cohesion is the density of a network, which is the number of existing lines in the network expressed as a proportion of the maximum number of possible lines. Table 6-5 shows the subnetwork densities for each biotechnology and nanotechnology cluster. It is evident that for networks of smaller sizes the density is higher and vice versa. Even though density is an indicator often used in social network analysis, it is more suitable to compare networks of the similar sizes, since density is inversely related to network size. De Nooy *et al.* (2005) explain that this is because the number of possible lines increases rapidly with the number of vertices, whereas the number of social ties, which each person can maintain is limited. Therefore the density was measured by the average degree of a network. The degree of a vertex is the number of lines that are directly connected to the vertex (Wasserman and Faust, 1994). It represents the number of direct collaborators with whom an inventor cooperated on at least one patent. The more co-inventors the inventors have, the tighter is the network structure. The average degree of a network then denotes the average of the degrees of all vertices and in fact it also shows the average number of co-inventors in each subnetwork, which was discussed earlier.

Accordingly, the biotechnology innovation subnetworks in the clusters of Saskatoon (average degree of subnetwork of 3.52), Edmonton (3.18) and Ottawa (3.06) are the densest and Montreal (2.94) and Vancouver (2.76) are still relatively dense. The innovation subnetwork in the nanotechnology cluster of Toronto is by far the densest in both Canadian biotechnology and nanotechnology (average degree of subnetwork of 5.32). The nanotechnology inventors in Toronto have direct or indirect access to a larger amount of information and a greater number of inventors than in any other cluster. Consequently the possibility for two inventors to get in touch through a chain of personal acquaintances is higher as well. Other larger nanotechnology clusters have much lower average degree values (Montreal 2.23, Ottawa 2.44 or Vancouver 2.8).

Since the biotechnology network is older, contains more inventors and is hence more developed it was expected to find it to be also denser, while that the connections between the subjects in the nanotechnology network were assumed to be much looser. However, this was not confirmed, mainly because of the very high cohesion among the nanotechnology inventors in the Toronto subnetwork. Moreover, also the average degree value for the full Canadian network shows a higher value (5) for nanotech than for biotech (4.26).

Table 6-5: Structural cohesion and cliquishness in biotechnology and nanotechnology cluster-based subnetworks

	Biotechnology			***Nanotechnology***		
	Structural cohesion		***Cliquishness***	***Structural cohesion***		***Cliquishness***
Cluster-based subnetwork	*Subnetwork density*	*Average degree*	*Average egocentric density*	*Subnetwork density*	*Average degree*	*Average egocentric density*
Toronto	0.003	2.42	0.44	0.011	5.32	0.54
Montreal	0.004	2.94	0.56	0.012	2.23	0.49
Vancouver	0.007	2.76	0.57	0.020	2.80	0.64
Ottawa	0.014	3.06	0.59	0.014	2.44	0.56
Edmonton	0.015	3.18	0.55	0.036	2.84	0.59
Calgary	0.022	2.00	0.29	0.078	2.48	0.45
Quebec	0.019	3.06	0.55	0.049	2.26	0.52
Kingston	0.022	2.04	0.47	0.061	2.06	0.45
Saskatoon	0.024	3.52	0.64	-	-	-
Winnipeg	0.018	1.40	0.32	-	-	-
Halifax	0.038	1.21	0.24	-	-	-
Sherbrooke	0.031	0.77	0.23	-	-	-
Network	***0.001***	***4.26***	***0.71***	***0.003***	***5.00***	***0.76***

6.3.4 Cliquishness in the subnetworks

Cliquishness is a property of a local network structure which refers to the likelihood that two vertices that are connected to a specific third vertex are also connected to one another. Cliquish networks have a tendency towards dense local neighbourhoods, in which individual inventors are better interconnected with each other. Such networks exhibit a high transmission capacity, since a great amount of knowledge could be diffused rapidly (Burt, 2001). Moreover, a high degree of cliquishness in an innovation network supports friendship and trust-building, and hence facilitates collaboration between innovators. Uzzi and Spiro (2005) and Schilling and Phelps (2007) argue that higher cliquishness enhances system performance and knowledge diffusion. However, Cowan and Jonard (2003) point out the existence of negative effects of cliquishness stemming from the loss due to repetition, as the knowledge exchanged in highly cliquish neighbourhoods is often redundant. Moreover, empirical findings of Fleming *et al.* (2006) confirm the negative impact of the higher cliquishness in the network on innovative productivity. The role of a high degree of cliquishness in the innovation production is still not obvious and the optimal degree will apparently depend on a variety of factors.

In this thesis the degree of local cliquishness for each vertex was measured with the egocentric density of a vertex, which is the fraction of all pairs of the immediate neighbours of a vertex that are also directly connected to each other, and then the average egocentric density of a subnetwork was calculated. The results are presented in Table 6-5. Cliquishness is quite comparable among the larger biotechnology subnetworks (Saskatoon, Ottawa, Vancouver and Montreal) or larger nanotechnology subnetworks (Vancouver, Edmonton, Ottawa and Toronto). The subnetworks of the smaller sizes in both networks seem to be less cliquish.

These results are, however, not in agreement with Newman (2001a) who found that the degree of network cliquishness in biomedicine is much lower than in other fields (clustering coefficient of 0.066), which he explained by the differences in social organization between biomedical and other research communities. The values for his other databases correspond to the results obtained in this thesis. The differences are probably caused by the distinct kinds of studied networks (as mentioned before, he created his networks based on the co-authorship of the scientific articles and not the co-inventorship of patents).

Both biotechnology and nanotechnology show quite comparable results regarding the cliquishness of the full networks.

6.3.5 Centrality of vertices

The centrality of a vertex indicates whether the position of an individual inventor within the subnetwork is more central or more peripheral. Inventors that are more central have better access to knowledge and better opportunities to spread information. Moreover, it is expected that inventors who occupy the most central positions in the subnetworks will be the most influential and probably the most prolific (star scientists). Two indicators of the vertex centrality were used in this thesis: degree centrality and betweenness centrality.

The simplest definition of centrality is the degree centrality of a vertex, which is in fact equal to the degree of the vertex defined above. Inventors in more central positions in the subnetwork are those directly connected to more other inventors and thus have more sources of knowledge at their disposal. Table 6-6 below shows the maximal centralities in all the subnetworks. The most connected inventors for both networks live in Toronto (the biotechnology one has 51 co-inventors while the nanotechnology one 42 co-inventors), but Montreal's most connected inventor has only 16 (biotech) or 15 (nanotech) direct collaborators. Other well connected inventors are located in Vancouver (27 biotech co-inventors and 16 nanotech co-inventors) and among the biotechnology subnetworks also in Saskatoon (25 co-inventors).

Betweenness centrality of a vertex is defined as a proportion of all shortest distances between pairs of other vertices that include this vertex (de Nooy *et al.*, 2004). An inventor is more central if a lot of shortest paths between pairs of other inventors in the subnetwork have to go through him. Betweenness centrality is therefore based on the inventor's importance to other inventors as an intermediary and it measures his control over the interactions between other inventors and thus over the flow of knowledge in the subnetwork. The highest betweenness centrality values come from the biotechnology inventors in Saskatoon (0.074) or Ottawa (0.068), whereas they are much lower for the nanotechnology inventors where the highest is in Edmonton (0.02).

In sum, the Saskatoon's most central biotechnology inventor occupies the most central location based on all three centrality measures. In nanotechnology, Toronto benefits from several quite central inventors (surpassing others, particularly in degree centrality), but so does Edmonton, with its most central inventor enjoying high maximum centrality levels as well.

Table 6-6: Centrality and centralization in biotechnology and nanotechnology cluster-based subnetworks

	Biotechnology				***Nanotechnology***			
Cluster-based subnetwork	*Vertex centrality*		*Subnetwork centralization*		*Vertex centrality*		*Subnetwork centralization*	
	Max degree	*Max betweenness*	*Degree*	*Betweenness*	*Max degree*	*Max betweenness*	*Degree*	*Betweenness*
Toronto	51	0.008	0.05	0.008	42	0.009	0.08	0.009
Montreal	16	0.011	0.02	0.011	15	0.005	0.07	0.005
Vancouver	27	0.005	0.06	0.005	16	0.013	0.09	0.013
Ottawa	16	0.070	0.06	0.068	9	0.008	0.04	0.007
Edmonton	20	0.019	0.08	0.019	14	0.021	0.15	0.020
Calgary	12	0.011	0.11	0.011	7	0.000	0.15	0.000
Quebec	8	0.003	0.04	0.003	5	0.008	0.06	0.008
Kingston	6	0.003	0.04	0.003	4	0.005	0.06	0.005
Saskatoon	25	0.076	0.15	0.074	-	-	-	-
Winnipeg	6	0.002	0.06	0.002	-	-	-	-
Halifax	5	0.010	0.13	0.010	-	-	-	-
Sherbrooke	2	0.000	0.05	0.000	-	-	-	-
Network	***66***	***0.009***	***0.01***	***0.009***	***54***	***0.006***	***0.02***	***0.006***

6.3.6 Centralization of the subnetworks

Contrary to centrality, which refers to positions of individual inventors, centralization characterizes an entire network. A highly centralized network has a clear boundary between the center and the periphery. The center of a centralized network allows more efficient transmission of knowledge, which consequently spreads fairly easily in highly centralized networks. A network is hence more centralized if centralities of the vertices vary substantially. Centralization of a network is defined as the variation in the degree centrality of vertices, divided by the maximum degree variation which is possible in a network of the same size (de Nooy *et al.*, 2004). Similarly as with centrality, two main measures of network centralization were used: degree centralization and betweenness centralization.

Degree centralization of a network is based on the variation in degree centrality of vertices in a network. The Saskatoon, Halifax and Calgary subnetworks show the highest degree centralization scores for the biotechnology clusters, while the Edmonton and Calgary subnetworks show the highest degree centralization scores for the nanotechnology clusters, which correspond to the same concepts.

Analogous to degree centralization, betweenness centralization of a network is based on the variation in betweenness centrality of vertices in the network. The results are shown in Table 6-6. It is again Saskatoon and Ottawa, which previously showed the highest maximal betweenness centralities of the vertices and now score the highest in betweenness centralization of all the biotechnology subnetworks as well. Among the nanotechnology clusters, Edmonton with the highest maximal betweenness centralities of the vertices has also high values for betweenness centralization.

In general, the biotechnology network has more highly central inventors than the nanotechnology network. For the centralization measures, the degree centralization indicator favours the nanotechnology network, whereas the betweenness centralization indicators show the reverse. As betweenness centralization refers to the positions of its inventors as intermediaries, it was not expected that the nanotechnology network would score higher because of its already mentioned disciplinary fragmentation. And indeed, the highest value is obtained in Edmonton, where the National institute for nanotechnology is located since 2001, with a much smaller score than Ottawa and Saskatoon which host similar National Institutes of the National Research Council of Canada for Biotechnology.

6.3.7 Geodesic distances in the subnetworks

A shortest path between two vertices is referred to as geodesic. The geodesic distance is then the length of a geodesic between them, which depends on the number of intermediaries needed for an inventor to reach another inventor in the subnetwork. A short path length in innovation networks should improve knowledge production and knowledge diffusion (Cowan and Jonard, 2004; Fleming *et al.*, 2004), since knowledge can move to the different parts of a network more quickly and spread rapidly among inventors. Moreover, as Cowan and Jonard (2004) suggest, decreased path length will cause knowledge to degrade less by bringing new sources of ideas and perspectives from farthest parts of the network to the inventors.

The longest geodesic in a network (the longest shortest path) is called the diameter of a network. It quantifies how much apart are the two farthest vertices in a network and it is a rough indicator of the effectiveness of a network in connecting pairs of inventors. In general, the observed diameters in the subnetworks seem to be fairly long when compared to the overall size of the components (see Table 6-7). This suggests a quite low connectedness in the subnetworks. The largest diameter among the biotechnology subnetworks is found in the Montreal and Ottawa clusters, where an inventor transmitting the knowledge needs as many as 10 intermediaries. In nanotechnology, it is also Ottawa which has the longest diameter. The exchange of knowledge is much easier in Vancouver (both among biotechnology and nanotechnology inventors) and obviously also in many other smaller clusters.

An indicator of the average distance of a network, which denotes an average of all the distances of all the vertices in the subnetwork, is a more global measure of efficiency in communication. Nevertheless, the distance between two unconnected vertices is not defined (does not exist) and the average distance hence could be measured only in fully interconnected networks. The average distance were therefore calculated only between reachable vertices (i.e., directly or indirectly connected). This measure shows similar results as the subnetwork diameter. The largest average distances are again found in the Montreal (4.27) and Ottawa (4.95) biotechnology clusters, and in Ottawa nanotechnology cluster (2.58). Obviously, the geodesic distances are also lower in smaller clusters. Moreover, it should be taken into consideration that the fact that the distances are calculated only between reachable vertices may bring a certain bias to these results, since any small or highly disconnected subnetwork should yield lower scores for geodesic distances. Therefore it is necessary to evaluate this measure more globally – while considering how many inventors could be reached within the cluster.

Table 6-7: Geodesic distances in biotechnology and nanotechnology cluster-based subnetworks

Cluster-based subnetwork	***Biotechnology***			***Nanotechnology***		
	Subnetwork diameter	*Avg distance (reachable vertices)*	*Max reach*	*Subnetwork diameter*	*Avg distance (reachable vertices)*	*Max reach*
Toronto	9	3.26	97	7	2.56	154
Montreal	11	4.27	108	4	1.67	20
Vancouver	5	1.98	37	3	1.61	20
Ottawa	11	4.95	74	10	2.58	22
Edmonton	7	2.50	48	3	1.60	14
Calgary	5	1.75	14	1	1.00	7
Quebec	4	1.48	10	3	1.22	5
Kingston	3	1.31	7	2	1.08	4
Saskatoon	6	2.75	53	-	-	-
Winnipeg	3	1.25	6	-	-	-
Halifax	2	1.23	5	-	-	-
Sherbrooke	1	1.00	2	-	-	-
Network	***17***	***6.55***	***578***	***17***	***4.16***	***335***

The reach of a vertex is defined as the number of vertices that can be reached from this particular vertex. Table 6-7 shows the maximal reach for each subnetwork, i.e. the maximum number of reachable inventors within a subnetwork. Evidently, more inventors could be directly or indirectly reached in larger networks. In the Montreal biotechnology subnetwork 108 inventors can reach each other, while in the larger Toronto biotechnology cluster it is only 97 inventors who are connected among themselves. In nanotechnology, however, the maximal reach in the Toronto cluster is 154. The clusters with shorter maximal reach are likely to be more disconnected and thus show lower scores of geodesic distances, whereas the clusters with highest numbers of reachable vertices are more connected and should show longer geodesic distances. The exception among the biotechnology subnetworks seems to be Saskatoon, which with a relatively long maximal reach does not show a very long average shortest distance. The Toronto nanotechnology subnetwork has a maximal reach many times longer than that of other nanotechnology clusters, but the average shortest distance of the Toronto subnetwork is not considerably longer than that of the other clusters and in fact even slightly shorter than that of the much smaller cluster of Ottawa. The geodesic characteristics of the Saskatoon biotech subnetwork and Toronto nanotech subnetwork are thus indicative of network structures which enable more efficient knowledge diffusion.

As expected, the biotechnology subnetworks show longer geodesic distances but also a much longer maximal reach. This is also evident in full Canadian networks (biotech has the average distance of 6.55 and maximal reach of 578 inventors, whereas nanotech has the distance of 4.16, but maximal reach of only 335 inventors). Knowledge should thus flow faster in nanotechnology cluster sub-networks.

6.3.8 Summary of the network properties

It was observed that in order to enhance the efficiency of each network in terms of knowledge diffusion, the network should be cohesive (which means that inventors are closely interconnected), cliquish (which fosters trust and close collaboration), it should have a long reach within large components (which enables bringing fresh and non-redundant knowledge from distant locations) and it should have a centralized structure (which supports fast knowledge transmission).

In biotechnology, the closest to these properties is the Saskatoon subnetwork. It is the densest, most cliquish and most centralized of all Canadian biotechnology clusters. It has on average the largest components and lowest share of isolates of all clusters. Despite the great size of the components, the diameter is still only of an average size. Inventors from both the Ottawa and Edmonton biotechnology clusters also benefit greatly from quite large components and fairly dense, relatively cliquish and rather centralised biotechnology subnetworks. The long geodesic distances however make it more difficult to bring new knowledge fast to all researchers. In contrast, the structural properties of the subnetworks of Calgary, Quebec and Toronto were not found to be very suggestive of efficient knowledge transmission and innovation generation. Both the Calgary and Quebec subnetworks are quite sparse and consist of the components of rather small sizes, suggesting great disconnectedness among inventors. Calgary, however, is quite centralized, which supports a more efficient transmission of knowledge, but it has a high share of researchers working in geographical isolation. Quebec is fairly cliquish and hence better interconnected. In both clusters, relatively short geodesic distances increase the speed of the knowledge transmission. The biotechnology subnetwork of the Toronto cluster is rather sparse, neither very cliquish nor centralized, and comprises components of relatively small sizes, many of which are completely isolate inventors. The Montreal biotechnology cluster, on the other hand, contains relatively large components through which knowledge has to travel large distances. It is also denser and more cliquish than the Toronto one; researchers seem to be more interconnected and knowledge could still be diffused more rapidly. The subnetwork structure of the Vancouver biotechnology cluster is somewhere in between the two previous patterns. It is denser than the Toronto subnetwork and quite cliquish, but comprises smaller components and thus involves shorter geodesic distances.

The most efficient nanotechnology cluster-based subnetworks are found in Toronto, Edmonton and Vancouver. Toronto's is the densest network of the Canadian nanotechnology clusters, where researchers are better interconnected and knowledge can hence be diffused quite rapidly. It has on average the largest components and the lowest share of geographically isolated researchers of all the clusters. Despite the great mean size of the components, the path lengths are still only slightly higher than average. Information can thus spread through a great number of researchers in a timely manner. The Toronto nanotechnology subnetwork is however only moderately cliquish and centralized. In contrast, inventors from both Edmonton and Vancouver nanotechnology clusters benefit from fairly cliquish and rather centralized nanotechnology subnetworks, the structure which supports both the trust-building among the researchers and a more efficient transmission of information through the centrally located researchers. The larger-sized components with quite short geodesic distances make it easier to bring new information fast to a relatively high number of inventors in both clusters. As for the nanotechnology clusters of Montreal and Ottawa, the structural properties of their subnetworks were found not to be very supportive of efficient knowledge diffusion and innovation generation. Both subnetworks are quite sparse and neither very cliquish nor centralized. They consist of the components of rather small sizes, which explains the relatively short path lengths measured in the networks. Also, a high percentage of researchers in both nanotechnology clusters work in a geographical isolation. These characteristics suggest a great disconnectedness among the inventors in a cluster.

6.4 Conclusions

The biotechnology and nanotechnology patenting in Canada has followed distinct paths. Base year for the start of biotechnology innovation in Canada could be considered the year of 1976, but the patenting really significantly accelerated only after 1987.

However, following the year 2001 Canadian biotechnology patenting has started to decrease. In case of nanotechnology, which is at an earlier stage of the industry life cycle, it is the year of 1986 that could be considered as its base year, after which the Canadian nanotechnology patent productivity has been almost always increasing. The annual patent growths in 2004 have thus reached almost comparable levels for biotechnology and nanotechnology.

Innovative activity in Canada is concentrated in several locations which roughly correspond to the larger metropolitan areas: 12 biotechnology and 8 nanotechnology clusters have been identified. In biotechnology, more than half of all Canadian inventors reside in three largest clusters - Toronto, Montreal and Vancouver, but in nanotechnology it is mainly the Toronto cluster which dominates the industrial sector since around one quarter of all Canadian inventors reside there. However, most of the innovations created by the Toronto nanotechnology inventors are owned by foreign assignees (mainly the US companies). Almost half of all the innovations authored or co-authored by Canadian nanotechnology inventors are assigned to the foreign subjects. Although Canadians do the research, the fruit of their labour is not appropriated within Canada. Canada therefore appears as a research subcontractor of patents in nanotechnology. This is not conducive to the creation of a healthy interconnected network of inventors where multidisciplinarity and diversity fosters invention.

The collaborative behaviour of the biotechnology and nanotechnology inventors inside and outside their clusters was also investigated. The majority of the all of these collaborations take place within the biotechnology or nanotechnology clusters and over a quarter are distant ties directed abroad. Most of the foreign collaborative ties are again linked to American inventors. Only a relatively small part of all the collaborations involves cooperation among the inventors from different clusters or with the out-of-cluster Canadians.

The intra-cluster collaborations within the cluster-based subnetworks were examined more in depth. The several structural network properties corresponding to each cluster were measured and then related to the likely efficiency of each subnetwork in the knowledge diffusion and the innovation creation. Moreover, a comparative analysis of the properties of the full networks and cluster-based subnetworks of biotechnology and nanotechnology was carried out and it was found that the collaborative structure within each technology to be quite distinct. The biotechnology innovation network is larger and more developed than the nanotechnology one. The biotechnology network was discovered to be also less fragmented. The specialization fields within the biotechnology are quite close in their scientific nature and are often overlapping. The inventors in the biotechnology network are thus more interconnected between each other. Nanotechnology, on the other hand, includes many quite disparate fields, where the inventors understandably work in a larger number of separated groups. A notable exception here is the Toronto nanotechnology cluster, which involves highly interconnected inventors with quite close collaboration ties and a dense subnetwork structure. The geodesic distances in biotechnology network are longer, but so is the maximal reach, which enables bringing fresh and non-redundant knowledge from distant locations. The cliquishness of both networks is however quite comparable, but its exact role in knowledge creation and innovation generation still remains to be determined.

The National Research Council of Canada has five national biotechnology institutes throughout the country, but only one in nanotechnology, a field much more fragmented. This analysis clearly shows that biotechnology cluster-based subnetworks are better developed and organised in a number of clusters in Canada, and especially in those hosting the five NRC institutes (Montreal, Vancouver, Ottawa, Saskatoon and Halifax although the latter is a much smaller cluster). These institutional effects have a positive influence on the organisation of innovation in these clusters. In contrast, in nanotechnology, two poles are present, Toronto and Edmonton, the latter still emerging. Around the National Institute for Nanotechnology, institutions are put in place to insure that discoveries are spun off or licensed to local firms in priorities in order to generate the synergies to the evolution of a successful cluster. Although the majority of the innovation capability lies in Toronto, it was observed that the majority of the intellectual property leaves the country. Other nanotechnology clusters are emerging, but their local network of inventors is still fragmented.

CHAPTER 7

GATEKEEPERS OF CANADIAN BIOTECHNOLOGY CLUSTERS

The last two chapters still remain within the realm of Canadian innovation, but the focus is shifted from the clusters and network structures to the individual inventors. The two chapters concentrate on the key individuals in the innovation process. While Chapter 8 explores the role and the network positions of the most prolific biotechnology inventors (star inventors), Chapter 7 studies the gatekeepers – the biotechnology inventors who occupy the key places in the innovation network, which allows them to fulfil the role of the suppliers of fresh information originating outside their own cluster. This chapter presents the way to identify them and to determine their relative importance as procurers of external knowledge for the cluster or for Canada.

7.1 Geographical and cognitive proximity

Wink (2008) proposes that gatekeepers can provide interface nodes between regional innovation systems by different forms of proximity. The ability of an actor to function as a gatekeeper thus depends on the kind of proximity which is necessary to span the boundary between the systems. There are several dimensions of proximity described in the literature. Torre and Gilly (2000) make a distinction between two different dimensions: geographic proximity, which refers to the spatial context, and organizational proximity, which is based on the organizational interaction of firms participating in clusters (and includes a cognitive dimension as well). Kirat and Lung (1999) incorporate the third dimension, institutional proximity, indicating the closeness among the agents influenced and restricted by the institutional environment. Boschma (2005) extends the classification and identifies five dimensions of proximity – cognitive (proximity related to the knowledge base of the actors), organizational (closeness of actors in organizational terms), social (closeness based on the socially embedded relations between agents, which involve trust, friendship, kinship and experience), institutional (proximity related to the institutional environment) and geographical (defined as the spatial or physical distance between economic actors). This chapter will focus mainly on the dimensions which are most relevant to the identification of the gatekeepers – geographical and cognitive proximities.

It is well established in economic geography to view regions as key drivers of innovation. This is built on the fact that geographical proximity facilitates knowledge sharing, since knowledge does not spill over large distances (Audretsch and Feldman, 1996; Jaffe and Trajtenberg, 1996). It is assumed that all firms in the cluster can benefit from these localized knowledge spillovers, which are not available to the firms outside the clusters. As a consequence, the firms in clusters are found to be more innovative (Baptista and Swann, 1998; Beaudry, 2001, Beaudry and Breschi, 2003). However, Boschma (2005) suggests that this view overemphasizes the role of geographical proximity in the transfer of knowledge between firms. He argues that other dimensions of proximity should be taken into consideration as well, since geographical proximity per se is neither a necessary nor a sufficient condition for learning to take place.

Another stream of literature on knowledge creation and diffusion emphasizes the role of cognitive proximity. As it was already discussed in section 1.2, some researchers (for example Cowan *et al.,* 2000; Breschi and Lissoni, 2001a, 2001b; Lissoni, 2001) argue that it is not geographic proximity which causes tacit knowledge to spill over between firms, but it is social connectedness of people in the network. Knowledge circulates and flows through the networks between the actors who are not necessarily placed in the same location. Technical or scientific knowledge is highly specific and its jargon differs from the jargon of the broader social community. The ones who understand it are the members of closed, restricted, but geographically dispersed "epistemic community", within which the tacit messages can be easily transmitted even if knowledge links take place among agents located far away in space. The networks thus do not require co-location of the actors for the production of innovation. On the other hand, physical proximity does not imply epistemic proximity, because epistemic communities are never as wide as to include all members of a local community. This means that firms in clusters may be excluded from knowledge sharing when they are not part of knowledge networks.

Apparently, the two concepts seem to stand against each other. Does it matter more for an inventor to be in the right location or to be connected to the right network of people? It has been argued that the combined effects of geographic and social spaces result in a more effective knowledge transfer (Owen-Smith and Powell, 2004; Sorenson and Stuart, 2001), while the causal relationship between the geographical and social distances has been suggested as well (Sorenson, 2003). In this thesis it is suggested that both the concept of space and the concept of network are at utmost importance for the knowledge creation and diffusion. Both geographical and cognitive dimensions nurture the growth of the cluster and promote innovation through a dynamic interaction of the actors localized in clusters who absorb external knowledge through the local and non-local networks. In order to bring the new knowledge to the cluster the gatekeepers thus have to be well connected both inside and outside the clusters.

This chapter explores the network architecture of Canadian biotechnology patenting and its role in knowledge transmission while considering two different collaboration spaces – geographical and technological. The geographical space in this context is based on the importance of geographic proximity and characterized by co-location of biotechnology firms in clusters. It assumes that knowledge networks are geographically localized and that no significant out-of-cluster linkages exist. Indeed, it has been shown in Chapter 5 that majority of all collaborative activities in Canadian biotechnology are carried out within clusters. Chapter 6 then revealed some of the collaboration characteristics of the geographical space in both biotechnology and nanotechnology networks. The technological space is understood here as the field in which collaboration ties are formed and knowledge is exchanged, while fully disregarding geographical aspects. It is assumed that all inventors who have collaborated on biotechnology innovations with each other at some point and are thus directly or indirectly interconnected in a network component are also part of the same epistemic community. Since the epistemic communities are restricted by the scientific fields and technological specializations this collaboration environment is in this thesis called a technological space.

In this chapter, the network architectures of geographical and technological collaboration spaces are compared and discussed and the level and nature of the overlap between them investigated. Finally, the points of interaction between the two collaboration spaces are explored and their importance for the cluster highlighted. These are the gatekeepers - the inventors who bridge over the geographical and technological collaboration spaces and thus enable the nurturing of biotechnology clusters with fresh knowledge originating outside.

7.2 Collaboration in geographical space

Based on the residences of inventors 12 Canadian biotechnology clusters have been identified in Chapter 4. It was shown that only a very small portion of inventors (around 3%) residing in Canada live outside the defined clusters and around 29% of inventors in this sample reside outside the Canadian borders.

Knowledge spillovers have already been discussed in the context of biotechnology innovation in Chapter 5. It was highlighted that the fact that biotechnology knowledge is largely tacit limits knowledge diffusion over long distances. As the transmission of tacit information and knowledge spillovers is usually associated with face-to-face contact, the collaboration among inventors working in clusters is thus encouraged by the benefits of acquiring the knowledge which the subjects located in close geographical proximity spill over. This section analyzes local collaborations carried out entirely within clusters, and as such the Canadian biotechnology innovation network is divided into geographically bound cluster subnetworks (as in Chapter 6). Each subnetwork strictly includes inventors who reside in that particular cluster, while excluding the ones that do not. The aim is to study how knowledge is transferred through these subnetworks.

Table 7-1 presents some of the main structural properties of the subnetworks created in this manner and is in fact a summary of the results obtained in Chapter 6. Table 7-1 shows that the cluster-based subnetworks are rather fragmented. Even though collaboration within clusters generally involves a very short geographical distance (commuting distance), inventors often choose to work in isolated groups. The fact that the largest components contain only 9%-18% of all inventors in each cluster confirms that inventors collocated within the same cluster are not highly interconnected. Furthermore, a substantial part of the cooperative links is directed outside the cluster. In Chapter 5 it was shown that Canadian inventors frequently take part in joint research projects including collaborators from abroad (29% of collaborations) or their colleagues residing in other clusters (11% of collaborations). The following section therefore disregards the geographical aspects and focuses solely on the technological space of collaboration.

Table 7-1: Structural properties of the cluster-based subnetworks (this is a summary of the results obtained in Chapter 6)

Cluster[1]	TRT	MTL	VAN	EDM	CAL	SAS	WIN	KIN	OTT	QUE	HAL	SHE	Canada
Number of inventors	927	698	411	210	91	147	77	94	224	127	33	26	4569
Number of patents[2]	834	466	255	153	127	98	33	63	279	57	20	16	2485[3]
COLLABORATION CHARACTERISTICS													
Number of collaborating pairs	1120	1027	568	334	91	259	54	96	343	155	20	10	9731
% of repeated collaborations	43%	36%	37%	37%	41%	28%	19%	33%	36%	18%	50%	20%	36%
Max number of repeated collaborations	60	11	10	14	16	8	3	10	19	7	5	3	60
STRUCTURAL COHESION													
Subnetwork density	0,003	0,004	0,007	0,015	0,022	0,024	0,018	0,022	0,014	0,019	0,038	0,031	0,001
Average degree	2,42	2,94	2,76	3,18	2,00	3,52	1,40	2,04	3,06	2,44	1,21	0,77	4,26
CENTRALIZATION OF SUBNETWORK													
Degree centralization	0,05	0,02	0,06	0,08	0,11	0,15	0,06	0,04	0,06	0,04	0,13	0,05	0,01
Betweenness centralization	0,008	0,011	0,005	0,019	0,011	0,074	0,002	0,003	0,068	0,003	0,010	0,000	0,009
CENTRALITY OF VERTICES													
Max degree centrality	51	16	27	20	12	25	6	6	16	8	5	2	66
Max betweenness centrality	0,008	0,011	0,005	0,019	0,011	0,076	0,002	0,003	0,070	0,003	0,010	0,000	0,009
GEODESIC DISTANCES													
Subnetwork diameter	9	11	5	7	5	6	3	3	11	4	2	1	17
Max reach	97	108	37	48	14	53	6	7	74	10	5	2	578
CLIQUISHNESS													
Average egocentric density	0,44	0,56	0,57	0,55	0,29	0,64	0,32	0,47	0,59	0,55	0,24	0,23	0.71
FRAGMENTATION													
# of components	342	218	134	67	39	34	44	38	70	44	20	18	894
Size of the 1st (2nd)largest component	98 (35)	109 (34)	38 (18)	49 (21)	15 (9)	54 (36)	7 (6)	8 (6)	75 (11)	11 (9)	6 (4)	3 (3)	579 (185)
Size of the 1st largest as % of all	11%	16%	9%	23%	16%	37%	9%	9%	33%	9%	18%	12%	13%
Size of the 2nd largest as % of all	4%	5%	14%	10%	10%	24%	8%	6%	5%	7%	12%	12%	14%
Ratio 2nd/1st largest	0,36	0,31	0,47	0,43	0,60	0,67	0,86	0,75	0,15	0,82	0,67	1,00	0,32
Average component size	2,71	3,20	3,07	3,13	2,33	4,32	1,75	2,47	3,20	2,89	1,65	1,44	5,11
Share of comp. formed by 50% of inventors	13%	11%	15%	11%	18%	6%	25%	24%	9%	21%	30%	33%	10%
Isolates as % of inventors	19%	15%	16%	17%	24%	13%	36%	18%	17%	15%	39%	46%	4%

[1]TRT	...Toronto	MTL	...Montreal	VAN	...Vancouver	EDM	...Edmonton	CAL	...Calgary	SAS	...Saskatoon
WIN	...Winnipeg	KIN	...Kingston	OTT	...Ottawa	QUE	...Quebec	HAL	...Halifax	SHE	...Sherbrooke

[2] The numbers are based on the residence of the assignees and only the patents with at least one Canadian assignee are thus included

[3] Also includes the patents assigned outside the clusters or co-assigned to the several clusters at the same time

7.3 Collaboration in technological space

In the technological space, collaboration is based on network components. All inventors in a component are directly or indirectly interconnected and it is thus supposed that they all collectively contribute to the innovation process. The attachment of inventors to their local environment is considered as secondary and the innovation network is analyzed regardless of the inventors' place of residence.

Canadian biotechnology inventors are grouped into 894 components, which suggests that the network is quite fragmented and that inventors are not highly interconnected. In terms of the number of vertices, the largest component (Component C1) includes 579 inventors, the second one (Component C2) consists of 185 inventors and the third (Component C3), of 175 inventors. There are few large components (10% of components include around 50% of inventors); most of them however are relatively small. As a consequence, the average number of inventors in a component is also relatively small (5.11). This is attributable to the fact that around 22% of all the components (195 components) are isolates (a component that consists of a sole inventor who has not collaborated).

The structure and main characteristics of the 30 largest components in the network are shown in Table 7-2. It is obvious that most components consist of inventors residing in several distinct clusters. This is particularly true for the largest components, where inventors are geographically spread over the entire country and abroad (Components C1 or C2). Some components, however, clearly consist of a great majority of inventors of one cluster. For instance, Component C3 seems to incorporate inventors from five Canadian clusters, but a closer inspection shows that 112 out of 124 Canadian inventors of Component C3 come from Montreal. The largest Montreal's component has 109 inventors (see Table 7-1) implying that there are only 3 Montrealers, which would be disconnected from the component if no inventors from other clusters were included. Similarly, 75 inventors of the largest component of Ottawa collaborate in Component

C2, which includes a total of 77 Ottawa inventors. Most of the other clusters' largest components are contained within Component C1, which looks like a great collaboration field for the most connected Canadian researchers except those from Montreal and Ottawa. In the case of Ottawa, this may be caused by the federal research concentration of the National Research Council seated in the Canadian capital, but Montreal is quite surprisingly isolated from the largest Canadian collaboration group of Component C1.

Some components (C6, C7, C19, C23 or C28) present intra-cluster cooperation within Canada, but also include some foreign cooperation relations. In fact, all of these 30 largest components include at least one foreign collaborator. Some of these "international" components consist of a majority of foreign inventors with only one or two Canadians (Components C10 or C14). These are probably much larger foreign networks in which a few Canadian inventors participate. For instance, Component C10 is based on collaboration on one single patent and is composed of 24 inventors; out of which 23 are foreign and only one is Canadian. Understandably, these mostly foreign components also show very low ratios of patents per inventor.[32]

---

[32] By concentrating on inventors of Canadian patents some much larger North American or even worldwide network which might link (indirectly) some of the components obtained may be missed. Since the focus here is on Canadian cluster gatekeepers, this does not constitute an obstacle to this study.

Table 7-2: Main characteristics and composition of the 30 largest components in the Canadian biotechnology innovation network

Component #	*C1*	*C2*	*C3*	*C4*	*C5*	*C6*	*C7*	*C8*	*C9*	*C10*	*C11*	*C12*	*C13*	*C14*	*C15*
# of inventors	579	185	175	78	50	44	39	36	30	29	27	27	27	24	23
# of patents	606	155	139	70	32	70	31	50	30	6	12	15	65	1	12
Patents/inventor	1.05	0.84	0.79	0.9	0.64	1.59	0.79	1.39	1.00	0.21	0.44	0.56	2.41	0.04	0.52
Number of the component's inventors in each cluster															
Toronto	154	16	8	16		35		5	22	1		25	22		1
Montreal	13	2	112	35	4		34								8
Vancouver	55	10	2		38			1			11		1		
Edmonton	50	1						2	1					1	
Calgary	20							9			3				
Saskatoon	54	40							2						
Winnipeg	1			7											
Kingston	9														
Ottawa	17	77	1	6											1
Quebec	9	2	1	1											1
Halifax		1						1							
Sherbrooke	2	2													
out-of-cluster	25	5		4								1			
abroad	170	29	51	9	8	9	5	18	5	18	13	1	4	23	12

Table 7-2: Main characteristics and composition of the 30 largest components in the Canadian biotechnology innovation network - continued

Component #	**C16**	**C17**	**C18**	**C19**	**C20**	**C21**	**C22**	**C23**	**C24**	**C25**	**C26**	**C27**	**C28**	**C29**	**C30**
# of inventors	23	22	20	19	18	18	18	18	17	17	17	17	16	15	15
# of patents	16	10	10	37	7	14	10	8	13	7	8	12	22	5	8
Patents/inventor	0.7	0.45	0.5	1.95	0.39	0.78	0.56	0.44	0.76	0.41	0.47	0.76	1.38	0.33	0.53
Number of the component's inventors in each cluster															
Toronto	8		1	18	1	2	1	6				1	15		
Montreal	5	19	12			7	12				12				1
Vancouver					3	5				1					
Edmonton					1				10						
Calgary						1			1					2	
Winnipeg										1					
Kingston					4				1						5
Ottawa												7			
Quebec	1														
Halifax									1						
out-of-cluster	2		1						1	1					
abroad	7	3	6	1	9	3	5	12	3	14	5	9	1	13	9

Table 7-3: Structural properties of the component-based subnetworks

Component Number	*C1*	*C2*	*C3*	*C4*	*C5*	*C6*	*C7*	*C8*	*C9*	*C10*	*C11*	*C12*	*C13*	*C14*	*C15*
Number of inventors	579	175	185	78	50	44	39	36	29	30	27	27	27	24	23
Number of patents	606	155	139	70	32	70	31	50	30	6	12	15	65	1	12
COLLABORATION CHARACTERISTICS															
Number of collaborating pairs	2057	560	517	185	167	105	83	92	336	89	61	44	46	276	87
% of repeated collaborations	45%	40%	29%	38%	46%	43%	40%	82%	82%	48%	16%	16%	54%	100%	62%
Max number of repeated collaborations	60	11	19	9	8	9	3	12	6	6	2	2	32	1	5
STRUCTURAL COHESION															
Subnetwork density	0,01	0,04	0,03	0,06	0,14	0,11	0,11	0,15	0,83	0,20	0,17	0,13	0,13	1,00	0,34
Average degree	7,11	6,40	5,59	4,74	6,68	4,77	4,26	5,11	23,17	5,93	4,52	3,26	3,41	23,00	7,57
CENTRALIZATION OF SUBNETWORK															
Degree centralization	0,10	0,10	0,10	0,12	0,45	0,37	0,24	0,27	0,19	0,26	0,27	0,16	0,48	0,00	0,27
Closeness centralization	0,19	0,19	0,16	0,22	0,49	0,31	0,37	0,34	0,28	0,36	0,32	0,23	0,52	0,00	0,42
Betweenness centralization	0,55	0,36	0,49	0,43	0,53	0,53	0,54	0,47	0,01	0,48	0,45	0,51	0,77	0,00	0,57
CENTRALITY OF VERTICES															
Max degree centrality	66	24	24	14	28	20	13	14	28	13	11	7	15	23	13
Max closeness centrality	0,24	0,29	0,24	0,32	0,62	0,48	0,51	0,50	1,00	0,57	0,51	0,39	0,63	1,00	0,71
Max betweenness centrality	0,56	0,38	0,51	0,48	0,56	0,56	0,59	0,51	0,02	0,52	0,51	0,60	0,81	0,00	0,59
GEODESIC DISTANCES															
Subnetwork diameter	17	12	14	10	7	6	7	7	2	5	6	8	6	1	3
Average distance	7,09	5,31	6,58	4,78	2,75	3,10	3,18	3,10	1,17	2,57	2,90	3,65	2,68	1,00	1,98
Max reach	57	174	184	77	49	43	38	35	28	29	26	26	26	23	22
CLIQUISHNESS															
Average egocentric density	0,80	0,81	0,79	0,75	0,80	0,57	0,82	0,77	0,94	0,80	0,83	0,69	0,69	1,00	0,94

Table 7-3: continued

Component number	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27	C28	C29	C30
Number of inventors	23	22	20	19	18	18	18	18	17	17	17	17	16	15	15
Number of patents	16	10	10	37	7	14	10	8	13	7	8	12	22	5	8
COLLABORATION CHARACTERISTICS															
Number of collaborating pairs	53	43	46	46	41	37	39	38	40	53	54	62	41	40	48
% of repeated collaborations	34%	7%	9%	59%	7%	70%	23%	0%	18%	32%	67%	19%	24%	5%	88%
Max number of repeated collaborations	3	3	3	26	2	3	3	1	5	2	5	12	12	2	4
STRUCTURAL COHESION															
Subnetwork density	0,21	0,19	0,24	0,27	0,27	0,24	0,25	0,25	0,29	0,39	0,40	0,46	0,34	0,38	0,46
Average degree	4,61	3,91	4,60	4,84	4,56	4,11	4,33	4,22	4,71	6,24	6,35	7,29	5,13	5,33	6,40
CENTRALIZATION OF SUBNETWORK															
Degree centralization	0,32	0,42	0,67	0,63	0,49	0,26	0,57	0,85	0,59	0,34	0,54	0,62	0,37	0,71	0,30
Closeness centralization	0,36	0,33	0,73	0,68	0,57	0,27	0,65	0,91	0,68	0,51	0,60	0,74	0,39	0,83	0,49
Betweenness centralization	0,58	0,62	0,71	0,67	0,68	0,44	0,70	0,85	0,69	0,50	0,36	0,50	0,42	0,71	0,50
CENTRALITY OF VERTICES															
Max degree centrality	11	12	16	15	12	8	13	17	13	11	14	16	10	14	10
Max closeness centrality	0,58	0,53	0,86	0,86	0,74	0,49	0,81	1,00	0,84	0,76	0,89	1,00	0,65	1,00	0,78
Max betweenness centrality	0,63	0,68	0,73	0,69	0,71	0,53	0,72	0,85	0,71	0,53	0,38	0,51	0,48	0,71	0,53
GEODESIC DISTANCES															
Subnetwork diameter	4	5	3	3	5	7	3	2	4	3	3	2	4	2	3
Average distance	2,49	2,76	1,96	1,89	2,17	2,89	2,01	1,75	1,95	1,93	1,68	1,54	2,20	1,62	1,85
Max reach	22	21	19	18	17	17	17	17	16	16	16	16	15	14	14
CLIQUISHNESS															
Average egocentric density	0,82	0,82	0,84	0,83	0,75	0,68	0,83	0,90	0,88	0,85	0,85	0,94	0,66	0,95	0,91

The network characteristics of the 30 largest network components could be found in Table 7-3. Four largest components usually show higher cohesion and lower centralization than smaller components. They obviously also have larger geodesic distances but higher maximal reach, since it takes longer for the information to travel all over the large component but it can reach many more other inventors. Striking exceptions to this pattern are two medium-sized components, in which all inventors (Component C14) or almost all inventors (Component C10) are connected to each other, since they have all collaborated with each other on all their patents (or almost all for Component C10). The larger components may however consist of several smaller components connected by a few individuals.

A comparison of the structural properties with the cluster-based subnetworks (Table 7-1) reveals that the component-based subnetworks (Table 7-3) are denser, more centralized and present more cliquishness, but they also have greater diameters. This should not be surprising as the cluster-based subnetworks are in facts smaller parts of components separated by the cluster of residence of its inventors. Collaboration within components is thus probably more efficient because higher structural cohesion of subnetworks indicates closer interconnectedness of inventors, higher cliquishness fosters trust and close collaboration, and higher centralization supports fast information transmission. In contrast, the cluster-based subnetworks show smaller diameters due to the high structural fragmentation. This means that the paths are shorter and information can travel faster in cluster-based subnetworks, but because of the smaller maximal reach, the information will finally be acquired by much less inventors. It is not unexpected that the transmission of knowledge through the network is more efficient if there are no geographical barriers and all the interconnected inventors could freely and frequently cooperate regardless of the distance between them. In reality, however, this is not usually the case. Even though collaboration of Canadian inventors with non-local partners is very common in biotechnology, it was shown in Chapter 5 that for most inventors, in fact, local intra-cluster collaborative relations are more frequent. Biotechnology inventors in Canada do take the geographical distance into consideration when searching for partners. Consequently both the technological and geographical spaces are considered to be extremely important concepts and the final task is thus to seek the points of interaction between the two spaces.

Since the cluster-based subnetworks consist of the local fragments (geographical space) of the component-based subnetworks (technological space), the aim is to find the key individuals who link both these spaces.

7.4 In a search of the gatekeepers

The last part of this chapter involves both cluster-based and component-based subnetworks and searches for the bridges between them. Here the objective is to understand exactly how the information travels among clusters through the component channels and to look for the inventors who bridge over the two spaces and thus enable the nurturing of biotechnology clusters with fresh external knowledge.

Every inventor can become a gatekeeper, but the ones who are very well interconnected both within and outside their clusters are best equipped and able to fulfil this function. In order to evaluate this ability, all Canadian inventors are first roughly categorized based on the nature of each inventor's connections with other inventors. Three categories of inventors are established: internal inventor, external inventor and intermediary. An internal inventor only has intra-cluster connections, i.e. no collaboration partner outside the cluster. An external inventor does not participate in any intra-cluster cooperation, since all of his links are directed out of the cluster. Even if he physically resides in the cluster he has no contacts there and any external knowledge which he acquires remains on the cluster's border. None of the internal or external inventors can thus contribute to the actual information transmission between clusters; an intermediary however maintains both intra-cluster and inter-cluster connections and as such, his existence is instrumental to delivering fresh outside knowledge to the cluster. Out of 3065 inventors residing in Canadian clusters, 31% (936 inventors) are such intermediaries. The remainder of this section evaluates these intermediaries in their role of procurers of nonlocal knowledge to the cluster.

The most obvious evaluation criterion is based on the amount of knowledge intermediaries bring into the cluster, which here corresponds to the number of direct sources of external knowledge to which each intermediary is connected. Table 7-4 shows the average number of inter-cluster links (or inter-lines, in the fourth column) for intermediaries in each cluster, which corresponds to the amount of potential knowledge an average intermediary delivers to his cluster. Moreover, the third column displays the average number of links (or average degree), including both intra-cluster (within the cluster) and inter-cluster (between clusters), that are connected to the intermediaries in each cluster. This measure indicates how well an average intermediary is interconnected in general. Furthermore, the intermediaries have been grouped based on the number of their inter-cluster links, the results of which are provided in the last four columns of the same table. Around 70% of all intermediaries collaborate with only 1 or 2 out-of-cluster partners and are thus connected to only 1 or 2 channels through which they can introduce external knowledge into the cluster. An intermediary with a low number of external connections could still be extremely important for the cluster as a transmitter of external information, since this also depends on his position in the network.

In order to evaluate the positions of the intermediaries in the network the notion of betweenness centrality was used. Since this measure does not distinguish between the place and direction of knowledge transmission (whether the inventor serves as an important intermediary mainly among the inventors from the same cluster or he is indeed instrumental in the external knowledge transfer to the inventors in the cluster), it cannot fully capture how strategic an inventor's position is as an external knowledge procurer.

Table 7-4: Inter-lines analysis for all intermediaries

	Number of gate-keepers	*as % of all*	*Average degree*	*Average number of interlines*	*Number of intermediaries with:*			
					1-2 inter-lines	*3-5 inter-lines*	*6-9 inter-lines*	*10 or more*
Toronto	247	27%	6.5	2.4	**187 (76%)**	44 (18%)	9 (4%)	7 (3%)
Montreal	244	35%	6.0	2.3	**174 (71%)**	53 (22%)	15 (6%)	2 (1%)
Vancouver	101	25%	5.7	2.2	**79 (78%)**	11 (11%)	6 (6%)	5 (5%)
Edmonton	92	44%	**7.3**	2.7	52 (57%)	27 (29%)	**13 (14%)**	
Calgary	35	38%	5.9	**3.3**	21 (60%)	7 (20%)	**5 (14%)**	**2 (6%)**
Saskatoon	36	24%	**8.8**	**3.2**	24 (67%)	4 (11%)	**6 (17%)**	**2 (6%)**
Winnipeg	27	35%	3.9	1.8	24 (89%)	3 (11%)		
Kingston	20	21%	4.8	2.4	13 (65%)	6 (30%)	1 (5%)	
Ottawa	97	43%	6.9	2.7	60 (62%)	27 (28%)	8 (8%)	2 (2%)
Quebec	29	23%	4.6	1.6	25 (86%)	4 (14%)		
Halifax	5	15%	5.0	1.8	3 (60%)	2 (40%)		
Sherbrooke	3	12%	3.3	2.0	2 (67%)	1 (33%)		
ALL	**936**	**31%**	**6.3**	**2.4**	**664 (71%)**	**189 (20%)**	**63 (7%)**	**20 (2%)**

At this point betweenness is thus used merely to filter out intermediaries whose betweenness is zero, since any external knowledge transmitted through such inventors is redundant. For instance, imagine an inventor *i* connected to the same exact inventors as at least one other inventor *j* in the component (who is a co-author on all the same patents as *i* and hence transmits exactly the same knowledge as the original inventor *i*). If inventor *j* has collaborated on a single additional patent without inventor *i*, then there is at least one other intermediary in the cluster which has exactly the same connections as the original inventor *i* plus at least one additional connection leading to other inventors. The obtained betweenness of the original inventor *i* will thus equal to zero. Betweenness in fact measures how the disappearance of an inventor would alter the shortest paths and connectedness between all other inventors. Since the disappearance of inventors with zero betweenness would neither reduce the amount of external knowledge which enters the cluster nor the speed at which it enters (no shortest path would get longer), they are considered redundant and hence excluded from further analysis. After this filtering process, only around half the intermediaries (434 or 14% of all Canadian inventors within clusters) are retained. Even though for the purpose of the analysis the redundant intermediaries are not further considered, they are nevertheless important in the regional system of innovation, as knowledge can "enter" the cluster from a number of sources. The reason to ignore these redundant gatekeepers for the moment will become apparent in the latter part of the chapter when the importance of such intermediaries as providers of outside knowledge to the cluster will be considered. Performing once again the interlines analysis exclusively for the non-redundant intermediaries yields Table 7-5 and allows a comparison with the previous results including all intermediaries (in Table 7-4).

Table 7-5: Inter-line analysis for non-redundant intermediaries only

	Number of gate-keepers	*As % of all*	*Average degree*	*Average number of interlines*	*Number of intermediaries with:* *1-2 inter-lines*	*3-5 inter-lines*	*6-9 inter-lines*	*10 and more*
Toronto	124	13%	9,0	3,2	**74 (60%)**	**36 (29%)**	8 (6%)	6 (5%)
Montreal	111	16%	8,1	3,0	**56 (50%)**	**42 (38%)**	11 (10%)	2 (2%)
Vancouver	39	9%	7,7	2,6	**26 (67%)**	**8 (21%)**	3 (8%)	2 (5%)
Edmonton	41	**20%**	9,9	3,1	17 (41%)	18 (44%)	6 (15%)	
Calgary	24	**26%**	7,1	**4,0**	12 (50%)	6 (25%)	**4 (17%)**	**2 (8%)**
Saskatoon	20	14%	**12,6**	**4,6**	9 (45%)	4 (2%)	**5 (25%)**	**2 (10%)**
Winnipeg	5	6%	4,6	1,8	4 (80%)	1 (2%)		
Kingston	10	11%	5,9	2,6	6 (60%)	3 (3%)	1 (10%)	
Ottawa	45	20%	9,5	3,4	21 (47%)	16 (36%)	6 (13%)	2 (4%)
Quebec	9	7%	7,1	1,9	6 (67%)	3 (33%)		
Halifax	3	9%	5,7	2,3	1 (33%)	2 (67%)		
Sherbrooke	3	12%	3,3	2,0	2 (67%)	1 (33%)		
ALL	**434**	**14%**	**8,6**	**3,2**	**234 (54%)**	**140 (32%)**	**44 (10%)**	**16 (4%)**

The comparison suggests that most redundant intermediaries have a very low number of ties to external knowlegde sources as the percentage of intermediaries with only 1 or 2 connections outside the cluster dropped from around 70% to about 50%. This shows that non-redundant intermediaries are usually better interconnected with out-of-cluster collaborators. A proportionally much greater amount of non-redundant intermediaries with many direct sources of external information (6 or more inter-lines) is found in the clusters of Saskatoon (35%) and Calgary (25%), whereas in the big clusters of Toronto, Montreal and Vancouver, almost 90% of all outside knowledge is brought into the clusters by less connected non-redundant intermediaries (1-5 inter-lines). In fact, this is already detectable in the analysis of all intermediaries in Table 7-4, but the exclusion of the redundant gatekeepers made this observation more pronounced. In Saskatoon and Calgary, gatekeepers have the highest average number of inter-lines. Furthermore, intermediaries from Saskatoon present the highest average degree for both redundant and non-redundant intermediaries. Intermediaries from these two cities therefore seem to be better interconnected with their external innovation environment than those in other clusters. These observations are however not surprising in the light of the integratedness of Saskatoon researchers within the two largest components identified in Table 7-2 - 94 out of the 147 inventors of Saskatoon collaborate within these two components. Very few inventors are present in the remaining 28 largest components (only two inventors in component C9).

Table 7-6 provides a list of the 25 non-redundant intermediaries with the highest number of direct sources of outside knowledge and orders them according to the number of their inter-cluster links. An inventor from Toronto (**$\mathbf{TRT_1}$**) has the highest number of direct external sources (29). The sum of the value of all his inter-lines is 81, i.e. the inventor has collaborated with 29 external collaborators on 81 occasions. The next column shows the degree of a vertex, which is the sum of all his links, including both inter-cluster and intra-cluster. The inventor **$\mathbf{TRT_1}$** has only four additional links within the cluster (his degree is 33), which means that all the external knowledge which he acquires flows further into the cluster only through 4 of his colleagues from the cluster. Since not all inventors in the clusters are interconnected within the cluster itself, it is not known how many of them benefit from the external knowledge introduced by any particular intermediary. These indicators do not allow the measurement of whether an inventor is alone in bringing external knowledge to these inventors or whether there are others contributing to this task (which would make his contribution less critical). Moreover, the amount of innovative potential this knowledge may create cannot be assessed. As a consequence, several measures to help answer these questions have been developed. In order to evaluate the importance of each inventor for the transmission of external knowledge and to assess the external innovative potential delivered by him to other inventors in the cluster a *Gatekeeper's Importance Index (GII)* both for the cluster and for Canada has been introduced.

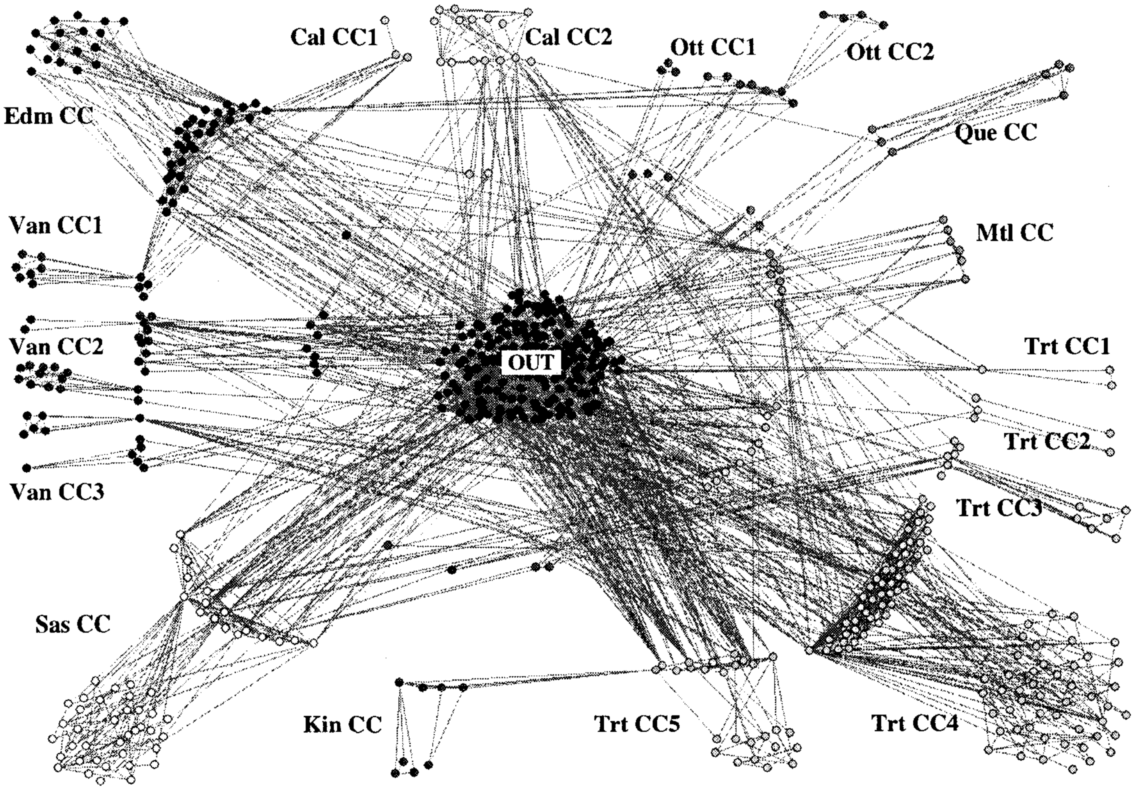
Table 7-6: First 25 non-redundant intermediaries with highest inter-clines count showing values of all importance indices and other network properties

Gate-keeper ID*	Rank by inter lines	Inter lines count	Inter lines value	Degree	Between-nes	Patents by the C-C	C-C size	Importance of the inventor for C-C	Importance of C-C for cluster	Importance of C-C for Canada	Gatekeeper`s Importance Index			
											($GII_i^{cluster}$)		(GII_i^{Canada})	
		(I_i)			($1000xB_i$)	(P_{cc})		(I_i/I_{cc})	($P_{cc}/P_{cluster}$)	(P_{cc}/P_{Canada})	Index value	Rank in cluster	Rank in Canada	Index value	Rank in Canada
TRT$_1$	*1*	29	81	33	1,63	55	25	24,37%	4,05%	1,55%	0,0161	***4th in TRT***	***12***	0,0062	***8***
TRT$_2$	*2*	21	54	28	2,71	55	25	17,65%	4,05%	1,55%	0,0194	***3rd in TRT***	***11***	0,0074	***6***
CAL$_1$	*3*	16	94	25	8,94	64	17	33,33%	30,48%	1,80%	0,9085	***1st in CAL***	***1***	0,0537	***1***
TRT$_3$	*4*	15	59	66	7,21	254	110	8,62%	18,72%	7,15%	0,1163	***1st in TRT***	***5***	0,0444	***2***
MTL$_1$	*5*	14	27	19	0,39	117	112	9,09%	15,70%	3,30%	0,0056	***1st in MTL***	***33***	0,0012	***21***
CAL$_2$	*6*	13	14	14	0,01	5	2	76,47%	2,38%	0,14%	0,0001	***11th in CAL***	***155***	0,0000	***195***
TRT$_4$	*7*	12	38	28	0,66	254	110	6,90%	18,72%	7,15%	0,0085	***8th in TRT***	***23***	0,0032	***14***
TRT$_5$	*7*	12	12	17	0,01	8	6	70,59%	0,59%	0,23%	0,0000	***49th in TRT***	***195***	0,0000	***154***
MTL$_2$	*9*	11	11	19	0,13	117	112	7,14%	15,70%	3,30%	0,0014	***9th in MTL***	***68***	0,0003	***51***
SAS$_1$	*9*	11	21	33	3,30	80	54	13,10%	51,28%	2,25%	0,2219	***2nd in SAS***	***3***	0,0097	***4***
SAS$_2$	*9*	11	16	36	5,13	80	54	13,10%	51,28%	2,25%	0,3443	***1st in SAS***	***2***	0,0151	***3***
TRT$_6$	*9*	11	23	17	5,30	55	25	9,24%	4,05%	1,55%	0,0199	***2nd in TRT***	***10***	0,0076	***5***
VAN$_1$	*9*	11	12	22	1,00	8	12	42,31%	2,00%	0,23%	0,0085	***2nd in VAN***	***22***	0,0010	***22***
OTT$_1$	*14*	10	15	14	0,10	18	6	43,48%	5,94%	0,51%	0,0025	***11th in OTT***	***52***	0,0002	***70***
OTT$_2$	*14*	10	59	16	0,01	13	7	34,48%	4,29%	0,37%	0,0001	***29th in OTT***	***169***	0,0000	***191***
VAN$_2$	*14*	10	25	15	0,27	7	6	100,00%	1,75%	0,20%	0,0048	***4th in VAN***	***35***	0,0005	***35***
CAL$_3$	*17*	9	14	21	2,06	64	17	18,75%	30,48%	1,80%	0,1179	***2nd CAL***	***4***	0,0070	***7***
KIN$_1$	*17*	9	9	12	0,01	4	4	56,25%	4,12%	0,11%	0,0002	***4th in KIN***	***129***	0,0000	***205***
MTL$_3$	*17*	9	13	24	0,38	117	112	5,84%	15,70%	3,30%	0,0035	***3rd in MTL***	***42***	0,0007	***30***
SAS$_3$	*17*	9	38	16	0,02	80	54	10,71%	51,28%	2,25%	0,0009	***9th in SAS***	***85***	0,0000	***123***
SAS$_4$	*17*	9	46	16	0,02	80	54	10,71%	51,28%	2,25%	0,0009	***11th in SAS***	***85***	0,0000	***125***
SAS$_5$	*17*	9	38	19	0,05	80	54	10,71%	51,28%	2,25%	0,0025	***6th in SAS***	***51***	0,0001	***85***
SAS$_6$	*17*	9	38	16	0,02	80	54	10,71%	51,28%	2,25%	0,0009	***10th in SAS***	***85***	0,0000	***123***
TRT$_7$	*17*	9	58	12	0,01	15	5	60,00%	1,11%	0,42%	0,0001	***43rd in TRT***	***176***	0,0000	***135***
TRT$_8$	*17*	9	9	17	0,44	254	110	5,17%	18,72%	7,15%	0,0043	***11th in TRT***	***39***	0,0016	***18***

[*] Gatekeeper ID is based on the cluster of the inventor's residence and his rank according to the number of inter-lines. (Whereas the ranking in the 12th column is based on the values of $GII_i^{cluster}$):

TRT$_{\#}$	...Inventor of rank # in Toronto	MTL$_{\#}$	...Inventor of rank # in Montreal	VAN$_{\#}$	...Inventor of rank # in Vancouver
KIN$_{\#}$	...Inventor of rank # in Kingston	CAL$_{\#}$	...Inventor of rank # in Calgary	SAS$_{\#}$	...Inventor of rank # in Saskatoon
OTT$_{\#}$	...Inventor of rank # in Ottawa				

First, here is the definition necessary for understanding the concept: A *Cluster-Component group of inventors (C-C group)* is a group of inventors residing in a Canadian cluster who are all directly or indirectly interconnected within the cluster. In a great majority of components, the C-C groups were created as a simple intersection between the clusters and the components, however - particularly in the 4 largest components - many inventors residing in the same cluster and being part of the same component are not directly connected within the cluster and end up in different C-C groups. Figure 7-1 illustrates the position of the three types of inventors of Component C1. In the centre of the figure is the largest group of inventors in this component, which is composed mainly of foreigners but also of some Canadian inventors residing outside clusters. It is fairly obvious that it is these predominantly foreign inventors who are interconnecting all other Canadian inventors in this component. Many of the inventors within the component do not have any other connection among themselves except through the foreign inventors. Canadian inventors residing in clusters are depicted here in three concentric circles around the core of foreigners and out-of-cluster inventors. The inner circle is composed of external inventors, which do not have any "direct" connections with their fellow inventors from the cluster, but indirectly through out-of-cluster and foreign inventors. Each of these external inventors actually constitutes a separate C-C group (those formed by the external inventors are neither indicated in the figure nor discussed further). In the middle circle are located the inventors connected to those residing both outside and inside the cluster – these are the intermediaries. The rest of the inventors - placed in the outer circle (on the periphery of the figure) - are internal cluster inventors connected only to intermediaries or among themselves. The C-C groups of Edmonton, Saskatoon and Kingston were created by the simple cluster-component intersection and there is thus only one C-C group for each cluster in this component. However, many inventors in other clusters had to be separated, notably in Toronto and Vancouver where they ended up in 5 different C-C groups in each cluster, since the only connections existing between them are through inventors residing outside clusters.



The vertices of different shades of grey indicate the inventors residing in different clusters.

Edm CC	...Edmonton C-C group	Mtl CC	...Montreal C-C group
Van CC#	...Vancouver C-C groups	Que CC	...Quebec C-C group
Sas CC	...Saskatoon C-C group	Ott CC#	...Ottawa C-C groups
Kin CC	...Kingston C-C group	Ca CC#	...Calgary C-C groups
Trt CC#	...Toronto C-C groups	OUT	...foreigners or Canadians outside clusters

Figure 7-1: Component C1 with all created C-C groups

The *Gatekeeper's Importance Indices (GIIs)* are based on the measurement of the importance of each intermediary as a source of external information for the C-C group to which he takes part and the importance of this C-C group either for the cluster or for Canada. The two *GIIs* are defined as:

$$GII_i^{cluster} = \frac{I_i}{I_{cc}} \times \frac{P_{cc}}{P_{cluster}} \times B_i \times 1000$$

$$GII_i^{Canada} = \frac{I_i}{I_{cc}} \times \frac{P_{cc}}{P_{Canada}} \times B_i \times 1000$$

where:

- $GII_i^{cluster}$... *Gatekeeper's Importance Index for Cluster* for inventor i
- GII_i^{Canada}... *Gatekeeper's Importance Index for Canada* for inventor i
- I_i...the number of inter-cluster links of the inventor i
- I_{CC}...the sum of all inter-cluster links of the C-C group cc (which includes inventor i)
- P_{cc}...the sum of all the patents invented or co-invented by at least one inventor from the C-C group cc (which includes the inventor i)
- $P_{cluster}$...the sum of all the patents authored or co-authored by all the inventors in the cluster in which the inventor i resides
- P_{Canada}... the sum of all the patents authored or co-authored by all the inventors residing in Canadian clusters
- B_i...betweenness centrality of the inventor i

The first term of the product in both indices captures the importance of the inventor as a source of external information for the C-C group. It measures the number of inter-links connected to each inventor (I_i) as a share of all the inter-links entering the given C-C group of inventors (I_{CC}). Since time is disregarded in this analysis and it is thus assumed that all links are active simultaneously, it can also be assumed that the amount of external knowledge incoming by each such channel is equal whatever the values of the links. The values of the links might show the efficiency with which the information is exchanged but do not reveal anything about the total amount of information which could be transmitted through the particular channel. This remains to be the same no matter how many times the collaboration between the two inventors took place and depends solely on the availability of the knowledge sources of the inventor on the other side of the channel. The second term of $GII_i^{cluster}$ evaluates the importance of each C-C group for the cluster based on the innovative productivity of that group. The patents which are authored or co-authored by at least one of the C-C group inventors are added for each group and divided by the sum of all the patents invented or co-invented by at least one of the inventors from the cluster ($P_{cluster}$). The last importance measure, which constitutes the second term of GII_i^{Canada} evaluates the importance of each C-C group for Canada and is based on the innovative productivity of the group as well. It also counts the number of patents which have been created within the C-C group of a given inventor and expresses that number as a share of the total innovative production in all Canadian clusters (P_{Canada}). The last term of the product in both indices measures the betweenness of the inventor (B_i) and indicates how well the inventor is interconnected in general[33]. This involves an overall evaluation of his network position which goes far beyond the external channels: it takes into consideration his other connections inside the cluster, the connections of all the inventors to whom he is connected and the positions of all the other inventors in the component from which he can indirectly gather knowledge or to whom he can deliver it. The resulting products are called *Gatekeeper's Importance Indices* and measure an inventor's importance as a procurer of external knowledge for the cluster ($GII_i^{cluster}$) or for Canada (GII_i^{Canada}) based on the share of innovative production to which he thereby contributes.

Table 7-6, which presents the importance measures for 25 intermediaries with the highest number of direct external sources, contains all the importance indices as well. Here are few examples which show how to interpret the measures: inventor **TRT$_1$** has the greatest count of inter-cluster collaboration links and contributes to around 24% of all the potential external knowledge input flowing into his C-C group (i.e. the percentage of TRT$_1$ interlinks with respect to the total number of interlinks of the cluster). The C-C group's share of the patent production represents around 4% of the cluster's production and around 1.5% of the total Canadian patent production. The final *Gatekeeper's Importance Indices*, which also take into account his network position, place inventor **TRT$_1$** in 8th position for his importance in the cluster and in 12th position for his importance in Canada. Within his own Toronto cluster, he is the 4th most important

---

[33] It is in part for the calculation of these indices that the redundant gatekeepers are ignored.

inventor in terms of his function as an intermediary of external information. Inventor **CAL_2** brings over 76% of external knowledge into the C-C group; this group however does not contribute significantly to the overall patent production in the cluster (2.4%) and even less in Canada (0.1%). Furthermore, even though **CAL_2** has 13 direct sources of information outside the cluster his C-C group inside the cluster is actually formed only by him and one additional inventor and his betweenness score is very low. In spite of the high number of external sources to which he has a direct access, the importance of such intermediary is quite negligible and he ranks very low both in his cluster and in Canada. Similar situation can be observed for the inventors **TRT_1**, **OTT_2**, **KIN_1** and **TRT_7**. These intermediaries utilize relatively many direct sources of external information for themselves, but they do not transfer the knowledge to many fellow inventors inside their own clusters. It would seem that these gatekeepers act in fact as ambassadors of knowledge from their own clusters to the outside world.

Four inventors with the highest scores of $GII_i^{cluster}$ in Canada are from the Saskatoon and Calgary clusters, which points out towards the crucial role played by these intermediaries in their own cluster. Table 7-7 presents the average importance indices for all inventors acting as intermediaries for the cluster. It shows that the average scores of $GII_i^{cluster}$ for Calgary (0.04) and Saskatoon (0.03) are much higher than that of any other cluster. The situation changes slightly when the average importance indices for Canada (GII_i^{Canada}) are calculated. Inventors from Toronto significantly gain in importance as gatekeepers for Canada (10 out of the first 20 intermediaries with the highest GII_i^{Canada} are from Toronto). Moreover, Table 7-7 shows that the average scores of GII_i^{Canada} are highest in Saskatoon (0.0026), Calgary (0.0013) and Toronto (0.0008), while it is much lower for other intermediaries.

Table 7-7: Average values of the indices of importance for the gatekeepers in each cluster

Cluster	*Importance of intermediary for the C-C*	*Importance of the C-C for cluster*	***Gatekeeper's importance index for cluster***	*Importance of the C-C for Canada*	***Gatekeeper's importance index for Canada***
Toronto	29.45%	6.16%	0.00204	**2.35%**	**0.00078**
Montreal	28.30%	5.01%	0.00030	1.05%	0.00006
Vancouver	39.61%	3.01%	0.00097	0.34%	0.00011
Edmonton	15.35%	18.47%	0.00510	1.36%	0.00037
Calgary	**39.57%**	9.74%	**0.04334**	0.58%	**0.00256**
Saskatoon	17.57%	**27.63%**	**0.02993**	1.21%	**0.00132**
Winnipeg	43.33%	6.49%	0.00046	0.14%	0.00001
Kingston	51.91%	4.95%	0.00029	0.14%	0.00001
Ottawa	11.51%	**29.94%**	0.00148	**2.56%**	0.00013
Quebec	34.55%	7.97%	0.00267	0.28%	0.00009
Halifax	58.33%	15.74%	0.00004	0.16%	0.00000
Sherbrooke	100.00%	12.64%	0.00003	0.10%	0.00000
Average	**28.36%**	**10.52%**	**0.00522**	**1.47%**	**0.00050**

Figure 7-2 displays both the absolute numbers and relative proportions of inventors in each cluster allocated to the categories of inventors based on their importance as procurers of external knowledge for the cluster, $GII_i^{cluster}$. Internal and external inventors do not participate in the transmission of external knowledge to the cluster, since they lack either the connection outside or inside their cluster. These inventors constitute the majority of inventors in all clusters (60%-80% for most clusters). Inventors which do maintain both intra-cluster and inter-cluster collaborations, but do not serve as indispensible intermediaries for other inventors are redundant intermediaries. As it was described, such intermediaries bring redundant external knowledge to the cluster, since not only would their disappearance not reduce the amount of external knowledge which enters the C-C group but it would not even make the shortest paths for that transmission longer. These inventors could still be productive and thus considered important creators of biotechnology innovation (even star scientists), but they are redundant as external information procurers. Around 15%-20% of inventors in most of the clusters are such intermediaries.

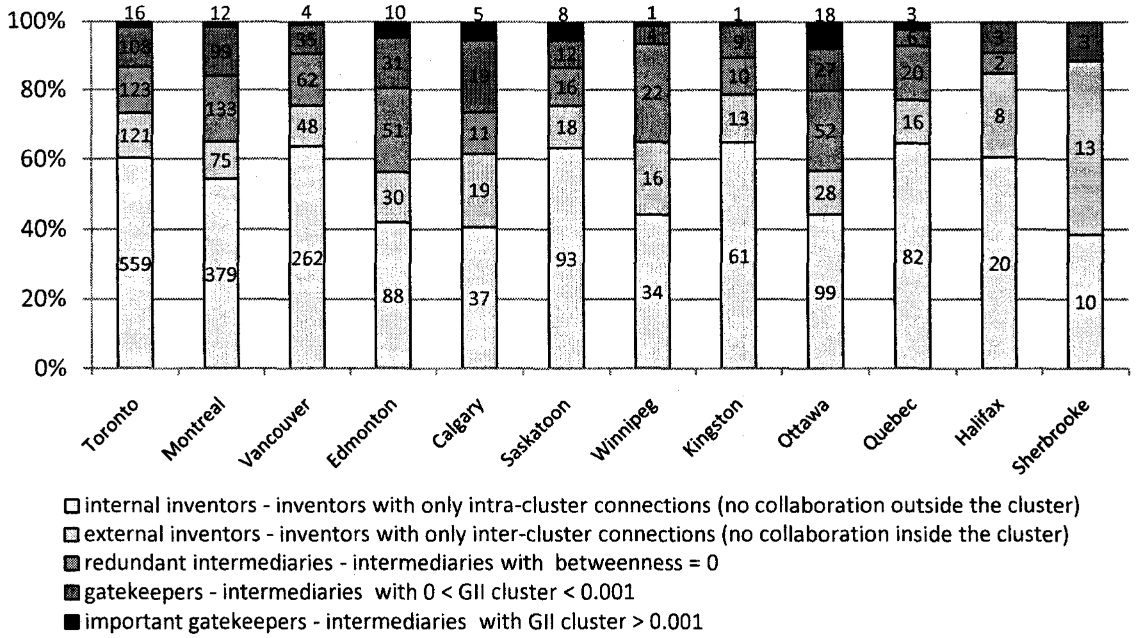


Figure 7-2: Numbers and relative proportions of inventors in the clusters categorized according to their importance as intermediaries

The remainder of the inventors are considered to be the gatekeepers. These are the intermediaries which do introduce non-redundant knowledge to the cluster and thereby contribute to the innovative potential of other inventors in the cluster. The highest percentage of gatekeepers among the cluster's inventors is found in Calgary (26%), Edmonton (20%) and Ottawa (20%), whereas Vancouver (9%) and the small clusters (6%-12%) have the lowest shares. However, the levels of contribution differ significantly among the gatekeepers themselves and therefore any gatekeeper with $GII_i^{cluster}$ of at least 0,001 has been designated as an important gatekeeper. Quite high percentages (around 60%) of all gatekeepers are considered to be important gatekeepers in the clusters of Saskatoon and Ottawa, but also in Quebec (30%), Edmonton (24%) and Calgary (20%). In the greatest clusters of Toronto, Montreal and Vancouver however, only around 10%-13% of all gatekeepers are important gatekeepers for the cluster (the number of the important gatekeepers in Ottawa is higher than their count in

Toronto even in absolute terms). There probably is a size effect here as the smallest clusters have none or very few of important gatekeepers.

Even though the analysis used in this research is quite static, in the real life the networks are often very dynamic and their structures keep changing. Consequently, the role of a gatekeeper is not permanent - the inventor may gain or lose some vital connections and thus change his importance as a procurer of external knowledge over time. Also, it is not known what would happen if a gatekeeper suddenly disappeared, so the study of the network dynamics would be interesting.

To briefly summarize the findings of this section concerning clusters: This analysis has shown that the proportions of gatekeepers among inventors is highest among the Calgary, Edmonton and Ottawa inventors. The clusters of Calgary and Saskatoon benefit from relatively many quite important and well-interconnected intermediaries with numerous direct sources of external knowledge and even from a couple of gatekeepers which are of extreme importance for the cluster's innovative productivity. In contrast, in the greater clusters of Toronto, Montreal and Vancouver the number of gatekeepers is proportionally lower; most of them are not of a very high importance for the cluster and also have a relatively low number of connections outside the cluster. The relative contribution of the Toronto inventors to the total Canadian biotechnology innovation production is however much more important.

Most of the network components (758 components, which represents 85% of all components) do not involve any gatekeeper. These are either components with only internal and external inventors (often single-inventor components or isolates) or components where all the inventors are connected to each other (each inventor is an intermediary who absorbs outside knowledge, but does not transmit it any further, since all of his colleagues have access to the same knowledge sources, i.e. they are all redundant intermediaries.). As for the components with gatekeepers (136 components, or 15% of the total), over half of them involve only one gatekeeper for the entire component. In this case there is one C-C group within the component where all external knowledge could be transferred to the group only through a single intermediary. If there are any other C-C groups within such component they consist either only of an external inventor or only of redundant intermediaries. Almost half (44%) of the 434 gatekeepers are part of the four largest components. This highlights the critical role played by the large components in the introduction of new knowledge to the cluster. Figure 7-1 illustrates the collaboration pattern among inventors within the largest component in the Canadian biotechnology network (Component C1, which involves 24% of all gatekeepers). It shows that inventors within the same cluster may not in fact be connected within the cluster and a foreign or out-of-cluster inventor is necessary to transmit knowledge between them. Within the same cluster and component there are groups working completely separately and the short geographical distance between them does not seem to play a role when seeking for collaboration partners. This allows making some conjectures about the position of the Canadian biotechnology network in the worldwide biotechnology innovation network. Many Canadian inventors who now seem to be disconnected may in fact be part of the same international component in the worldwide biotechnology innovation network. The complete Canadian biotechnology network would then be in fact much less fragmented than it can be seen now and there may exist one giant Canadian biotechnology network component, which would comprise a great majority of inventors as suggested by Newman (2001a). Furthermore, if this theory is extended further, most biotechnology inventors in the world might in fact be united in one giant international component where they all indirectly collaborate, share their knowledge and create collective inventions.

7.5 Conclusions

The geographical space and the technological space overlap to a certain extent, but differ in their structure. Many inventors from the same cluster may also be part of the same network component. The bulk of smaller components are entirely contained within one cluster, larger components however usually encompass several clusters. Moreover, most of the larger or medium-sized components include some foreign cooperation relations as well. These foreign inventors are extremely important in connecting Canadian inventors from different clusters together (or even from the same cluster - particularly in the largest components), which makes their presence critical for the transmission of knowledge between Canadian inventors. it was conjectured that if all biotechnology patents in the world were included in the analysis, the Canadian biotechnology network would be less fragmented and most of the inventors would in fact be a part of one giant international biotechnology innovation component in which all inventors indirectly collaborate, share their knowledge and create collective inventions.

The points of interaction between the geographical and technological spaces of collaboration have also been investigated. In order to understand exactly how knowledge travels among clusters through the channels of components, the search for gatekeepers was carried out. The gatekeepers are the inventors who bridge over the two spaces and thus enable the nurturing of biotechnology clusters with fresh external knowledge. In order to identify these gatekeepers, indicators, which measure each inventor's importance as a procurer of external knowledge for the cluster (or for Canada) based on the share of innovative production to which he thereby contributes, have been developed. Only around 10%-20% of all inventors in most clusters were identified as gatekeepers and are responsible for the inflow of external information to the cluster. Since the affiliations of the inventors are not known in this thesis it cannot be determined from which environment the gatekeepers arise – whether they are academics, industrial or governmental inventors. This would be an interesting topic for further investigation. However, some of the properties of the gatekeepers could be explored, especially those related to their patenting activity: Are the most productive inventors also the best procurers of external knowledge for the cluster? This is the question which will be answered in the next chapter which deals, among others, also with the intermediary role of star scientists.

CHAPTER 8

STAR SCIENTISTS IN CANADIAN BIOTECHNOLOGY NETWORK

The literature review in Chapter 1 has provided the details on research regarding the phenomenon of star scientists, which originates mainly in the United States. However, there are not many studies concentrated on Canada. Zucker and Darby (1996b) found no evidence of substantial star involvement (star affiliated with or linked to a biotechnology company) by Canadian biotechnology stars. Moreover, according to their results Canada was indicated as the major loser of key talent in biotechnology by migration (together with Switzerland and United Kingdom). The Canadian losses presumably reflect the ease of mobility to the particularly attractive US market.

According to Queenton and Niosi (2003), however, Canadian biotechnology clusters are strongly related to high-class academic research and especially to the star scientists working in universities. Their study also highlights the importance of geographical proximity of star scientists for obtaining the venture capital, and for starting and growing the biotechnology firm. It was confirmed that also in Canada many of the star scientists capitalise on their knowledge through firm start-ups. Niosi (2003) estimates that one third of Canadian biotechnology firms are university spin-offs.

This chapter adds to the research on the biotechnology star scientists in Canada. A new method of identification of the star scientists, which involves both the quantity and quality of the patents, is proposed. Moreover, a network approach is adopted and the positions of the star scientists in a complex net of innovative collaborations are examined.

8.1 Identification of the star scientists in Canadian biotechnology

Zucker and Darby (1996) created the definition of the biotechnology star scientists based on the number of genetic sequence discoveries or the number of articles reporting genetic sequence discoveries. Queenton and Niosi (2003), who searched for biotechnology stars in Canada, included the number of genetic sequence discoveries, the number of publications and the number of patents in their definitions. In this research patents are considered to be the main discriminatory indicator. The prominent researchers in the dataset are defined either based on patent quantity only, or based on both the quantity and quality simultaneously. Moreover, the examination of the most prominent researchers based on their record of forward citations in scientific articles was included as well.

First, using only the number of patents as a discriminatory indicator was considered. The numbers of patents authored or co-authored by each inventor are displayed in Figure 8-1. It is evident that most inventive output is produced by only a small percentage of the most prolific inventors some of which are listed in Table 8-1[34]. Every inventor with more than 15 patents is defined to be *prolific*; according to this classification, 51 prolific inventors in Canadian biotechnology are identified (which is around 1% of all inventors). Then among these, 22 inventors are considered to be *star* scientists, defined here as all the inventors with more than 20 patents. Four of the most prolific inventors have made a significantly greater contribution to the biotechnology innovation than other inventors and produced more than 50 patents. These individuals will be called *superstars*. As an example, the most productive inventor in Canadian biotechnology has registered 151 patents. This is considerably more than any other researcher in the group (see Figure 8-1), and may be caused by a "lab director effect".

---

[34] The list has been anonymised, the first letter represents the town of residence of the inventor and the digit subscript its rank as a prolific inventor.

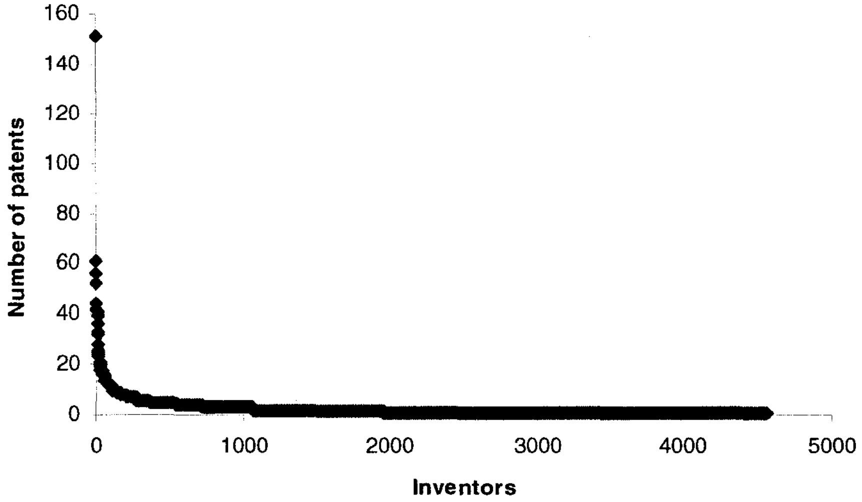


Figure 8-1: Distribution of the number of patents authored or co-authored by each inventor

During this research it has been observed that there are great differences in patent quality (measured again by the average number of the patent claims). You can see in Table 8-1 that the most prolific inventors do not necessarily register patents with the highest value. Therefore it was decided to incorporate patent quality as a second discriminatory factor when defining star scientists. A *Quantity and Quality Patent Index (QQ Index),* which takes into consideration both patent counts and the mean patent value for each inventor, has been created:

$$QQI_i = \frac{N_i * C_i^{avg}}{C^{avg}}$$

where

- QQI_i value of the QQ Index indicator for inventor i;
- N_i number of patents at the USPTO invented by inventor i;
- C_i^{avg} average number of patent claims for all patents at the USPTO by inventor i;
- C^{avg} average number of patent claims for all inventors in the database.

This indicator modifies the number of patents according to the gap between the average number of claims of a particular inventor and an average number of claims for all the inventors in the database. For an inventor which produces patents of an average value, this indicator should be equal to his number of patents, whereas the other inventors can improve or worsen their standing depending on the average quality of their patents. According to the *QQ Index* a *QQ-prolific* inventor is defined as one with a minimal *QQ Index* value of 20 (there are 50 of such inventors, which again represent around 1% of all the inventors) and a *QQ-star* inventor as one with an index greater than 30 (22 of such star inventors). Three inventors with the highest value of *QQ Index* are called *QQ-superstars*. Table 8-1 shows that for the most prolific inventors, the picture has not changed dramatically, but was slightly modified. Many inventors in the database however had to give up their prominent positions and, on the other hand, many have substantially improved their ranking.

The third indicator which was used to find the prominent inventors is related to the more scholarly side of a researcher's qualities. The number of forward citations to the researchers' articles represents a scientist's ability to contribute to knowledge development. *ISI Web of Knowledge*^{SM} provides a tool to identify individuals that have made fundamental contributions to the advancement of science and technology in recent decades. It lists the most highly cited individuals within several broad subject categories for the period 1981-1999 (later years are not currently available)[35]. The list includes only the researchers with a really extraordinary accomplishment, since it comprises less than 0.5% of all publishing researchers in the database. The data obtained from the list of highly cited scientists in biotechnology has been merged into the database of inventors. It has been found that 28 of the inventors are also highly influential scientists and scholars as illustrated in the last column of Table 8-1.

---

[35] As a consequence of the lack of the more current observations, older scientists with an extensive publication record probably have a better chance of being classified as star scientists because of the extensive observation period, whereas younger scientists who already belong to the very top of their class may still not be included because they have not yet accumulated enough publications and citations.

Table 8-1: Positions of all prominent inventors according to three different measures

*Inventor**	***Number of patents***			*Average # of claims*	***Quantity and Quality Patent Index***			*Highly cited scientist*
	Ranking	*Patent count*			*Ranking*	*QQI value*		
T_1	**1**	151	superstar	12	**1**	110,25	QQ-superstar	
T_2	**2**	61	superstar	12	**3**	44,54	QQ-superstar	
T_3	**3**	56	superstar	11	**11**	37,48	QQ-star	
T_4	**4**	52	superstar	10	**21**	31,64	QQ-star	
T_5	**5**	44	star	11	**24**	29,45	QQ-prolific	
M_1	**6**	42	star	14	**12**	35,78	QQ-star	
O_1	**7**	41	star	14	**13**	34,92	QQ-star	
S_1	**8**	39	star	16	**10**	37,97	QQ-star	
T_6	**9**	36	star	15	**19**	32,86	QQ-star	
T_7	**10**	33	star	12	**33**	24,09	QQ-prolific	
T_8	**10**	33	star	9	**65**	18,07		
T_9	**12**	32	star	13	**32**	25,31	QQ-prolific	
T_{10}	**13**	28	star	16	**28**	27,26	QQ-prolific	highly cited
C_1	**14**	26	star	10	**74**	15,82		
T_{11}	**14**	26	star	25	**7**	39,55	QQ-star	
T_{12}	**14**	26	star	12	**55**	18,98		
T_{13}	**17**	25	star	15	**40**	22,82	QQ-prolific	
T_{14}	**17**	25	star	15	**40**	22,82	QQ-prolific	
O_2	**19**	24	star	15	**43**	21,90	QQ-prolific	
S_2	**20**	23	star	24	**17**	33,59	QQ-star	highly cited
E_1	**21**	21	star	9	**136**	11,50		
T_{15}	**21**	21	star	9	**136**	11,50		
C_2	**23**	20	prolific	16	**53**	19,47		
T_{16}	**23**	20	prolific	11	**98**	13,39		
O_3	**25**	19	prolific	16	**60**	18,50		
O_4	**25**	19	prolific	12	**93**	13,87		
Q_1	**25**	19	prolific	16	**60**	18,50		highly cited
T_{17}	**25**	19	prolific	11	**108**	12,72		
C_3	**29**	18	prolific	25	**27**	27,38	QQ-prolific	
C_4	**29**	18	prolific	21	**37**	23,00	QQ-prolific	
E_2	**29**	18	prolific	18	**52**	19,71		highly cited
O_5	**29**	18	prolific	12	**102**	13,14		
T_{18}	**29**	18	prolific	31	**16**	33,95	QQ-star	
T_{19}	**29**	18	prolific	17	**58**	18,62		
T_{20}	**29**	18	prolific	17	**58**	18,62		
T_{21}	**29**	18	prolific	9	**173**	9,86		
T_{22}	**29**	18	prolific	9	**173**	9,86		
T_{23}	**29**	18	prolific	9	**173**	9,86		
T_{24}	**39**	17	prolific	16	**72**	16,55		
T_{25}	**39**	17	prolific	9	**200**	9,31		highly cited
T_{26}	**39**	17	prolific	8	**228**	8,27		

*Inventor**	***Number of patents***			*Average # of claims*	***Quantity and Quality Patent Index***			*Highly cited scientist*
	Ranking	*Patent count*			*Ranking*	*QQI value*		
V_1	**39**	17	prolific	7	**294**	7,24		highly cited
C_5	**43**	16	prolific	18	**66**	17,52		
C_6	**43**	16	prolific	12	**129**	11,68		
E_3	**43**	16	prolific	8	**253**	7,79		
E_4	**43**	16	prolific	6	**390**	5,84		
M_2	**43**	16	prolific	10	**177**	9,73		
O_6	**43**	16	prolific	14	**95**	13,63		
O_7	**43**	16	prolific	12	**129**	11,68		
T_{27}	**43**	16	prolific	19	**60**	18,50		
T_{28}	**43**	16	prolific	18	**66**	17,52		
T_{29}	**52**	15		22	**46**	20,08	QQ-prolific	
M_3	**57**	14		61	**2**	51,96	QQ-superstar	
V_2	**57**	14		28	**35**	23,85	QQ-prolific	
M_4	**67**	13		49	**9**	38,76	QQ-star	
O_8	**67**	13		44	**14**	34,80	QQ-star	
T_{30}	**67**	13		33	**29**	26,10	QQ-prolific	
T_{31}	**67**	13		33	**29**	26,10	QQ-prolific	
F_1	**79**	12		39	**25**	28,47	QQ-prolific	
T_{32}	**79**	12		38	**26**	27,74	QQ-prolific	
T_{33}	**79**	12		14	**163**	10,22		highly cited
M_5	**94**	11		36	**33**	24,09	QQ-prolific	highly cited
F_2	**130**	9		42	**37**	23,00	QQ-prolific	
M_6	**130**	9		75	**4**	41,07	QQ-star	
M_7	**130**	9		75	**4**	41,07	QQ-star	
M_8	**130**	9		74	**6**	40,52	QQ-star	
M_9	**130**	9		61	**18**	33,40	QQ-star	
O_9	**130**	9		42	**37**	23,00	QQ-prolific	
F_3	**158**	8		41	**47**	19,96	QQ-prolific	
F_4	**158**	8		41	**47**	19,96	QQ-prolific	
F_5	**158**	8		41	**47**	19,96	QQ-prolific	
M_{10}	**158**	8		80	**8**	38,94	QQ-star	
M_{11}	**158**	8		71	**15**	34,56	QQ-star	
M_{12}	**158**	8		64	**22**	31,15	QQ-star	
O_{10}	**158**	8		41	**47**	19,96	QQ-prolific	
T_{34}	**158**	8		46	**42**	22,39	QQ-prolific	
T_{35}	**158**	8		11	**451**	5,35		highly cited
M_{13}	**204**	7		77	**20**	32,79	QQ-star	
M_{14}	**204**	7		71	**23**	30,24	QQ-prolific	
O_{11}	**266**	6		60	**43**	21,90	QQ-prolific	
V_3	**266**	6		64	**36**	23,36	QQ-prolific	
W_1	**266**	6		9	**821**	3,29		highly cited
S_3	**361**	5		68	**45**	20,69	QQ-prolific	highly cited
T_{36}	**361**	5		7	**1368**	2,13		highly cited
V_4	**361**	5		84	**31**	25,55	QQ-prolific	

*Inventor**	*Number of patents*		*Average # of claims*	*Quantity and Quality Patent Index*		*Highly cited scientist*
	Ranking	*Patent count*		*Ranking*	*QQI value*	
T_{37}	**718**	3	12	**1270**	2,19	highly cited
T_{38}	**718**	3	6	**2479**	1,10	highly cited
out_{1}	**1067**	2	6	**3080**	0,73	highly cited
T_{39}	**1067**	2	9	**2479**	1,10	highly cited
V_{5}	**1067**	2	15	**1536**	1,83	highly cited
V_{6}	**1067**	2	11	**2067**	1,34	highly cited
V_{7}	**1067**	2	8	**2699**	0,97	highly cited
V_{8}	**1067**	2	6	**3080**	0,73	highly cited
M_{15}	**1956**	1	8	**3678**	0,49	highly cited
M_{16}	**1956**	1	4	**4184**	0,24	highly cited
T_{40}	**1956**	1	26	**1774**	1,58	highly cited
T_{41}	**1956**	1	17	**2607**	1,03	highly cited
T_{42}	**1956**	1	15	**2831**	0,91	highly cited
T_{43}	**1956**	1	10	**3402**	0,61	highly cited
T_{44}	**1956**	1	5	**4074**	0,30	highly cited
V_{9}	**1956**	1	38	**1238**	2,31	highly cited

* $T_{\#}$ ...	Toronto inventor	$C_{\#}$ ...	Calgary inventor	$Q_{\#}$ ...	Quebec inventor
$M_{\#}$ ...	Montreal inventor	$S_{\#}$ ...	Saskatoon inventor	$out_{\#}$ ...	inventor outside cluster
$V_{\#}$ ...	Vancouver inventor	$W_{\#}$ ...	Winnipeg inventor	$F_{\#}$...	foreign inventor
$E_{\#}$ ...	Edmonton inventor	$O_{\#}$ ...	Ottawa inventor		

Surprisingly, the three distinct indicators of the prominent inventors showed quite different results. This methodology has enabled to identify 101 prominent inventors (95 of them from Canadian clusters). Only two (T_{10} and S_{2}) scientists/inventors are however indicated as prominent by all three measures, 24 inventors are considered to be prominent by two of the indicators, 18 of which are concurrently identified by both the number of patents and the *QQ Index*. The two measures are obviously much more correlated together than with the indicator of highly cited scientists. For example, two of the five existing superstars are also *QQ-superstars* and two others are *QQ-stars*. The remaining 75 inventors were identified as prominent by only one measure. Among them, one *QQ-superstar* scientist falls to the 57th rank if the patent value is not taken into consideration and many other *QQ-prolific* inventors would occupy a rank as low as the 361st rank. The value of a patent hence seems to be an important discriminatory factor. Furthermore, some of the highly cited inventors reach even lower positions based on the other two indicators (as low as 1956th rank), since eight of the highly cited scientists have only one biotech patent at the USPTO, while six of them appear as inventors only on two patents. Moreover, the fact that only 28 matching scientists were found in both lists suggests that there must also be many highly influential biotechnology researchers (as acknowledged by their citing colleagues) who never filed any patent application at the USPTO[36]. These highly cited scientists are assumed to come mostly from an academic environment, where the publication performance is more appreciated and more rewarding than impressive patent scores. The scientists with the most prolific publication record may thus often neglect patent application opportunities. A much less probable explanation is that these highly cited scientists simply patent their inventions at different patent offices (e.g. CIPO or EPO). Table 8-2 presents the results per cluster by including all the discussed measures. Toronto is the leader in the number of prominent scientists in the cluster (44 scientists out of which 4 are superstars) while Montreal and Vancouver are far behind (16 and 9 scientists, respectively). In terms of sheer number of patents, Toronto excels (15 stars) but when the quality is taken into consideration, Montreal has in fact more *QQ-star scientists* who produce patents of high value than Toronto (11 stars compared to 7 stars). The Toronto cluster also houses the highest number of scientists with an outstanding citation record (12 scientists), whereas Montreal is lagging behind with only 3 highly cited scientists in the database. Vancouver's record is more modest on all fronts (it has virtually no star and only 3 prolific scientists), except the indicator of highly cited researchers (with 6 such scientists). The small clusters of Kingston, Halifax and Sherbrooke do not enjoy the benefits of any prominent scientist.

---

[36] The difference in the publication and patent records has been already observed at the Japanese corporate scientists in the pharmaceutical industry (Furukawa and Goto, 2006). The most frequently publishing scientists did not apply for a considerably greater number of patents than other researchers in their companies, but they had a positive effect on the number of patent applications filed by their co-authors.

Table 8-2: Prominent inventors by cluster

Biotechnology cluster	*Stars (superstars)*[a] *Number of patents*	*QQ Index*	*Prolific inventors*[b] *Number of patents*	*QQ Index*	*Highly cited scientists*	*All prominent researchers Total number*	*as % of all inventors*
Toronto	15 (4)*	7 (2)*	28	18	13	44 (4)*	**4.75%**
Montreal	1	11 (1)*	2	13	3	16 (1)*	2.29%
Vancouver			1	3	6	9	2.19%
Edmonton	1		4		1	4	1.90%
Calgary	1		6	2		6	**6.59%**
Saskatoon	2	2		3	2	3	2.04%
Winnipeg					1	1	1.30%
Ottawa	2	2	7	5		11	4.91%
Quebec			1		1	1	0.79%

[a] The numbers in brackets denote the number of these stars that are considered to be superstars in the cluster.

[b] All prolific inventors (including stars and superstars)

To conclude this section, it was found very fruitful to use the multi-indicator approach for the analysis of the prominent scientists. The picture became much more complete when the patent value was included in the equation instead of the sole patent count. Star scientists or highly prolific biotechnology inventors were found that they do not necessarily author or co-author patents of the highest value. By taking into consideration patent quality, the ranking of star and prolific inventors has changed. Not all prominent and highly cited researchers and scientists in biotechnology produce patents or register them at the USPTO, and for those that do, their patents are not of the highest value in terms of number of claims.

8.2 The positions of the star and QQ-star scientists in the network

This section investigates the positions of the stars and *QQ-stars* in the Canadian biotechnology collaboration network. Moreover, the level of overlapping of the prominent researchers and the gatekeepers is explored.

8.2.1 Hypotheses

It is expected to find evidence of the crucial role played by star scientists in biotechnology networks by occupying very central network positions. The central position in the network structure usually implies that star scientists are connected to a much larger number of inventors than others. Obviously, the identification of both the star or *QQ-star* scientists is directly related to the number of patents they produced and thus it can certainly be assumed that the scientists with higher number of patents or higher *QQ Index* score will have a higher number of collaborators. The first hypothesis thus reads as follows:

- *Hypothesis H_1a: The inventors with a higher patent production have more collaborators.*
- *Hypothesis H_1b: The inventors with a higher QQ Index have more collaborators.*

Since the inventors with a higher number of patents are usually more central they are also much better interconnected in the complex net of interrelationships. The central network position of the star or *QQ-star* scientists enables them to reach all other inventors in the network faster, because the length of the shortest paths between them and other inventors is usually greatly reduced due to the numerous connections they have. As a consequence, the stars and *QQ-stars* are able to get a much improved access to knowledge in the network. This is the core of the second hypothesis:

- *Hypothesis H_2a: The inventors with a higher patent production enjoy better access to information because of the reduced shortest paths to all the other inventors in the network.*
- *Hypothesis H_2b: The inventors with a higher QQ Index enjoy better access to information because of the reduced shortest paths to all the other inventors in the network.*

The star and *QQ-star* inventors are expected to also have more strategic positions in the network in terms of their ability to control the flow of information between other inventors. Their highly central positions enable them to act as intermediaries for the transfer of knowledge between many other inventors in the network. This increased flow of knowledge thus gives them a greater power over the knowledge distribution among others. The existence of the star and *QQ-star* inventors is thus crucial for a great number of other inventors in the network and their disappearance from the network would not only slow down the knowledge flow in the whole network by increasing the lengths of the shortest paths among many others, but it would also completely disconnect many inventors and thus highly limit their knowledge sources. Therefore the inventors with a higher number of patents or *QQ Index* scores are supposed to have more strategic network positions and thus have control over a greater flow of information, hence the third hypothesis:

- *Hypothesis H_3a: The inventors with a higher patent production have control over a greater amount of knowledge which passes through them.*
- *Hypothesis H_3b: The inventors with a higher QQ Index have control over a greater amount of knowledge which passes through them.*

The local neighbourhood of the star and *QQ-star* inventors is also expected to be more cliquish. The stars or *QQ-stars* have direct or indirect access to a larger number of other innovators and it is therefore assumed that their local environment will be also more dense and cohesive. This should support friendship and trust-building, and thus facilitate collaboration between the innovators. This is thus expected to also be a contributing factor to the success of the stars and *QQ-stars*, which leads to the fourth hypothesis:

- *Hypothesis H_4a: The inventors with a higher patent production are positioned in more cliquish local neighbourhoods.*
- *Hypothesis H_4b: The inventors with a higher QQ Index are positioned in more cliquish local neighbourhoods.*

Star and *QQ-star* inventors are expected to also play a crucial role in the nurturing of clusters with fresh knowledge originating outside. They have more collaborators, many of which probably reside in different clusters or even countries. The abundant connections outside their own clusters should enable the stars and *QQ-stars* to serve as knowledge gatekeepers - as procurers of external knowledge for other inventors which collaborate less or focus on joint research within the same region (for the more detailed discussion on gatekeepers see Chapter 7). The star and *QQ-star* inventors will thus be among the few inventors who are responsible for the inflow of external information to the cluster and will also play much more important role in nurturing of clusters with fresh outside knowledge, and hence the fifth hypothesis reads:

- *Hypothesis H_5a: The inventors with a higher patent production play more important role as gatekeepers for the clusters in which they reside.*
- *Hypothesis H_5b: The inventors with a higher QQ Index play more important role as gatekeepers for the clusters in which they reside.*

Similarly, the importance of the star and *QQ-star* inventors as procurers of external knowledge for Canada is assumed to also be much higher than the importance of less prolific inventors. Now the focus is on the more general impact of inventors which import the external knowledge to other Canadian inventors and the importance of that knowledge for Canada in terms of its contribution to the innovative potential. The sixth hypothesis therefore proposes that the inventors with higher number of patents or higher *QQ Index* will also play much more important role as procurers of outside knowledge for Canada:

- *Hypothesis H_6a: The inventors with a higher patent production play more important role as gatekeepers for Canada.*
- *Hypothesis H_6b: The inventors with a higher QQ Index play more important role as gatekeepers for Canada.*

8.2.2 Methodology and results

In order to validate the above hypotheses the most common measure of correlation – the Pearson Product Moment Correlation, which reflects the degree of linear relationship between two variables - is used. The correlation coefficients for each two variables for every hypothesis are calculated as explained in Table 8-3. Various indicators of the structural network properties (for their more detailed description see section 6.3 in Chapter 6) are used as variables representing the attributes of the inventors' positions: The number of collaborators of each inventor in H_1 is calculated as the degree centrality, which measures the number of lines that are connected to each vertex. The closeness of the inventors to all other inventors in the network in H_2 is measured by the closeness centrality of each vertex expressed as the number of other vertices divided by the sum of all distances between the vertex and all others. However, since closeness centrality is calculated only among the inventors who are directly or indirectly interconnected, it would be misleading to use this measure for all of them. The vertices in small components would show very high centralities, because all the inventors are close to each other, but it would not reveal much about their centrality in comparison with the inventors included in other components. The total correlation is then expected to be greatly underestimated. Therefore, only the inventors who are interconnected in the largest network component (579 inventors) are included in the testing for H_2. The amount of information which passes through each inventor in H_3 is calculated with the betweenness centrality of each vertex, which measures the proportion of all shortest distances between pairs of other vertices that include this vertex. Average egocentric density, which is a fraction of all pairs of the immediate neighbours of a vertex that are also directly connected to each other, is used to measure the degree of local cliquishness for each inventor in H_4. Finally, in hypotheses H_5 and H_6 the indices defined in the previous chapter (for the exact definitions see section 7.4 of Chapter 7) are used: *Gatekeeper's Importance Index for Cluster* ($GI^{cluster}$ *Index)* and *Gatekeeper's Importance Index for Canada* (GI^{Canada} *Index)*, each calculated for every vertex in the network.

The resulting values of the correlation coefficients are shown in Table 8-4. Since for large samples it is usually easy to achieve significance of the correlation, all of the correlation coefficients were found highly significant. The strength of the relationship was thus used to determine if the relationship explains very much or not. The variables were considered uncorrelated if $r < 0.1$, weakly correlated if $0.1 < r < 0.3$, moderately correlated if $0.3 < r < 0.5$ and strongly correlated if $0.5 < r < 1.00$.

Table 8-3: The examined variables for each hypothesis

$H_{\#}$	*Variable #1*	*Variable #2*	$H_{\#}$	*Variable #1*	*Variable #2*
H_1a	Number of patents	degree centrality	H_1b	*QQ Index*	degree centrality
H_2a		closeness centrality	H_2b		closeness centrality
H_3a		betweenness centrality	H_3b		betweenness centrality
H_4a		avg egocentric density	H_4b		avg egocentric density
H_5a		$GI^{cluster}$ *Index*	H_5b		$GI^{cluster}$ *Index*
H_6a		GI^{Canada} *Index*	H_6b		GI^{Canada} *Index*

Table 8-4: Pearson correlation coefficients

	Degree centrality	*Closeness centrality*	*Betweenness centrality*	*Avg egocentric density*	$GI^{cluster}$ *Index*	GI^{Canada} *Index*
# of patents	0.594**	0.391**	0.454**	-0.118**	0.237**	0.510**
QQ Index	0.550**	0.344**	0.344**	-0.085**	0.184**	0.370**

** Correlation is significant at the 0.01 level

The results for each hypothesis are summarized in Table 8-5. The correlations were found to be very strong in both hypotheses H_1a and H_1b. This was expected, since the number of patents of an inventor is usually related to the number of his collaborators. By every jointly created invention leading to a new patent the inventor usually also gains new collaborators (unless he continues to work always with the same group of inventors in all of his patents). Also the more central positions of the inventors with a higher number of patents (or higher *QQ Index* scores) were confirmed in both H_2a and H_2b and both H_3a and H_3b. The most important inventors in terms of patent counts and patent quality do play more important roles in the networks. They have a better ability to reach all the knowledge in the network due to the reduced length of the shortest paths to all other inventors. Moreover, their positions are highly strategic, since they enable them to assume control over a great flow of information.

Hypotheses H_4a and H_4b were however not confirmed. The inverse but extremely weak relationships between the number of patents or level of *QQ Index* and the level of the local cliquishness of an inventor were found to exist. This means that the assumption of dense and cohesive relationships supporting friendship and trust-building, thereby facilitating collaboration between the innovators and thus contributing to the innovativeness was not supported. The research carried out in the area of network cliquishness (already discussed in section 6.3.4 of Chapter 6) has also not been conclusive so far. On the one hand, Uzzi and Spiro (2004) and Schiling and Phelps (2007) show that high cliquishness in the networks enhances the system's innovative performance. On the other hand, the empirical findings of Fleming *et al.* (2006) prove the negative impact of the higher degree of cliquishness in the network on the innovative productivity. The authors argue that there is an optimal degree of cliquishness that depends on a variety of factors. Cowan and Jonard (2003) identify both positive and negative effects of high cliquishness on knowledge growth. They argue that the net effect is determined by both the benefits from differentiated neighbourhoods (agents in various neighbourhoods highly differ) and the loss due to repetition (cliquishness duplicates transmissions). The obtained very weak negative correlation supports the view that there may be negative effects of high cliquishness on knowledge growth. More empirical research is needed to clarify this relationship.

Finally, the two hypotheses related to the gatekeepers have mixed results. The importance of a gatekeeper as the procurer of external knowledge for the cluster in which he resides in H_5a and H_5b proved to be only very weakly correlated to the number of patents or *QQ Index*; however the importance of gatekeeper for Canada was found to be either strongly (H_6a) or moderately (H_6b) correlated. This is not surprising, since the number of patents or *QQ Index* are calculated globally as is the GI^{Canada} *Index* but $GI^{cluster}$ *Index* reflects only the local role of the intermediary in his own cluster. His role could be very significant in certain smaller or medium-sized clusters, but at the same time quite negligible in terms of his contribution to the overall Canadian innovative potential.

Table 8-5: The results for all the hypotheses

$H_{\#}$	*Correlation*	*H confirmed?*	$H_{\#}$	*Correlation*	*H confirmed?*
H_1a	strongly correlated	YES	H_1b	strongly correlated	YES
H_2a	moderately correlated	YES	H_2b	moderately correlated	YES
H_3a	moderately correlated	YES	H_3b	moderately correlated	YES
H_4a	inverse relationship	NO	H_4b	inverse relationship	NO
H_5a	weakly correlated	NO	H_5b	weakly correlated	NO
H_6a	strongly correlated	YES	H_6b	moderately correlated	YES

8.2.3 Are stars and QQ-stars also gatekeepers?

Even though both the gatekeepers and the star scientists are important and usually quite influential inventors, the two concepts are rather distinct. An inventor or scientist could be highly prolific (in terms of the number of patents and/or scholarly articles), but if he is not well connected he still might not bring any external knowledge into the cluster and thus cannot play the role of the gatekeeper. Similarly, a very important gatekeeper (who has vital connections leading to internal and external inventors who themselves have very good further connections) may happen not to be very productive in terms of patent applications or scholarly articles. Therefore it is essential to identify the level of overlap between the star or *QQ-star* inventors and the gatekeepers.

Since the correlation between the patent counts or *QQ Index* and the *Gatekeeper's Index of Importance for Canada* is confirmed, the aim is to see how many of the stars or *QQ-stars* are also gatekeepers of significant importance for Canada. Table 8-6 shows absolute and relative numbers of the star, *QQ-star* and highly cited scientists who belong to the five categories of inventors based on their network positions and the level of the *Gatekeeper's Index of importance for Canada,* as they were defined in the previous chapter. Internal and external inventors do not participate in the transmission of external knowledge to the cluster, since they lack either the connection outside their cluster (internal inventors) or inside their cluster (external inventors). These inventors constitute the majority of inventors in all the clusters. The stars, *QQ-star* or highly cited scientists are rarely external inventors, but around 14% of stars, 45% of *QQ-stars* and 22% of highly cited scientists are internal inventors who collaborate exclusively within their own cluster. Inventors which do maintain both intra-cluster and inter-cluster collaborations, but do not serve as indispensible intermediaries for other inventors were called redundant intermediaries. Such intermediaries bring redundant external knowledge to the cluster, since not only would their disappearance not reduce the amount of transmitted external knowledge but it would not even make the shortest path for that transmission longer. These inventors could be theoretically still quite productive and thus considered important creators of biotechnology innovation, but they are not essential as external knowledge procurers. As the results in Table 8-6 show, there are no stars or *QQ-stars* among the redundant intermediaries, but 22% of highly cited scientists belong to this inventor category. Gatekeepers are the intermediaries which do introduce non-redundant knowledge to the cluster and thereby contribute to the innovative potential of other inventors in Canada. The inventors with the top highest scores of GI^{Canada} *Index* were called here very important gatekeepers. The table shows that 86% of all star inventors are gatekeepers (27% are gatekeepers and 59% are very important gatekeepers), 55% of all *QQ-stars* are gatekeepers (23% are gatekeepers and 33% are very important gatekeepers) and 49% of all the highly cited scientists are gatekeepers as well (27% are gatekeepers and 22% are very important gatekeepers).

Thus it can be concluded that the majority of the star and *QQ-star* scientists are also gatekeepers. However, the relationship between the stars and the gatekeepers seems to be stronger than the relationship between the *QQ-stars* and the gatekeepers. This was expected since there the GI indices do not involve the patent quality. This was also confirmed by the correlation coefficients in the previous analysis, which were showing somewhat higher strength for the relationship between the number of patents and GI^{Canada} *Index* (Hypothesis H_6a) and also in all the cases of the vertex centralities (Hypotheses H_1a, H_2a and H_3a) than for the same relationships with the *QQ Index* (Hypotheses H_1b, H_2b, H_3b and H_6b). This suggests that even if the *QQ Index* may be a more accurate measure of the inventor's importance in terms of his inventive contribution as it reflects both the quantity and the quality of his patents, it is less accurate when assessing the importance of the position for the inventor in the network and his importance as gatekeeper. It is the number of patents but not the quality of these patents, which is related to the ability of the inventor to acquire external knowledge and to nurture the clusters with information from outside.

Table 8-6: Overlapping of stars, QQ-stars and highly cited scientists with gatekeepers

Stars	*Very important gatekeepers*	*Gatekeepers*	*Redundant intermediaries*	*Internal inventors*	*External inventors*	*Total*
Prolific only[a]	11 (38%)	6 (21%)	3 (10%)	9 (31%)		29 (100%)
Stars only[b]	9 (50%)	6 (33%)		3 (17%)		18 (100%)
Superstars	4 (100%)					4 (100%)
All prolific[d]	24 (47%)	12 (24%)	3 (6%)	12 (24%)	0 (0%)	51 (100%)
All stars[c]	**13 (59%)**	**6 (27%)**	0 (0%)	3 (14%)	0 (0%)	22 (100%)

QQ-stars	*Very important gatekeepers*	*Gatekeepers*	*Redundant intermediaries*	*Internal inventors*	*External inventors*	*Total*
QQ-prolific only[a]	6 (26%)	7 (30%)	2 (9%)	7 (30%)	1 (4%)	23 (100%)
Stars only[b]	5 (26%)	5 (26%)		9 (48%)		19 (100%)
QQ-superstars	2 (67%)			1 (33%)		3 (100%)
All QQ-prolific[d]	13 (29%)	12 (27%)	2 (4%)	17 (38%)	1 (2%)	45 (100%)
All QQ-stars[c]	**7 (32%)**	**5 (23%)**	0 (0%)	**10 (45%)**	0 (0%)	22 (100%)

Highly cited	*Very important gatekeepers*	*Gatekeepers*	*Redundant intermediaries*	*Internal inventors*	*External inventors*	*Total*
All highly cited	**6 (22%)**	**7 (27%)**	**6 (22%)**	**6 (22%)**	2 (7%)	27 (100%)

[a] Only inventors who are prolific but are not considered to be also stars or superstars
[b] Only stars who are not considered to be also superstars
[c] All prolific inventors (including stars and superstars)
[d] All stars (including superstars)

8.3 Conclusions

The first objective of this chapter was the identification of the prominent researchers in Canadian biotechnology clusters. It was proposed to take into consideration the patent quality when identifying the prolific inventors, and developed a measure which includes both the patent count and the patent value in the equation. Star scientists or highly prolific biotechnology inventors are observed not to necessarily author or co-author patents of the highest value. Furthermore, the scientists whose publications are the most highly cited have also been identified. The results show that not all the prominent researchers and superior scientists in biotechnology produce patents or register them at the USPTO. An explanation based on the differences in the reward systems in academic and industrial environments was offered.

In the second part the positions of the stars and *QQ-stars* in the network structure have been studied. The results show that the inventors with higher number of patents and the higher *QQ Index* assume more central positions in the network: they have more collaborators, they enjoy better access to information because of the reduced shortest paths to all the other inventors in the network and they also have greater control over the knowledge flows in the network since much more information passes through them. Nevertheless, the hypothesis regarding the level of cliquishness in the inventor's neighbourhood was not confirmed: The inventors with a higher number of patents or levels of *QQ Index* do not assume the network positions with a higher level of the local cliquishness. The impact of cliquishness of individual inventors on their innovative propensity has not been empirically studied so far and the existing research regarding the innovative performance of the networks with various degrees of cliquishness has not been conclusive. The very weak negative correlation however supports the view that there may be also negative effects of high cliquishness on knowledge growth as proposed by Cowan and Jonard (2003) and empirically supported by Fleming *et al.* (2006).

Finally, the relationship between the stars or *QQ-stars* and the gatekeepers was investigated. It was found that the great majority of star inventors (86%) and of QQ-star inventors (55%), and almost half of all the highly cited scientists (49%) are also gatekeepers responsible for the inflow of external knowledge which highly contributes to the Canadian innovative potential. However, it is only the number of patents but not the quality of these patents, which is related to the ability of the inventor to acquire external knowledge and to nurture the clusters with information from outside.

CHAPTER 9

CONCLUSIONS AND RECOMMENDATIONS

9.1 Concluding remarks

The common thread running throughout the thesis is the study of industrial clusters. The first issue examined concerns the role of the industrial composition of a cluster. The question how the composition of economic activities in the cluster influences the growth of the region has been asked by many researchers before who ended up with quite inconsistent answers. This thesis brought together a large range of studies, which have provided substantial academic support for the positive impact of both Marshall and Jacobs externalities on regional performance. In addition, a non negligible number of negative MAR effects imply that specialisation of a region may also hinder economic growth. Diversification is much less likely to produce this negative impact. This thesis has made a major contribution in clarifying why the results are often conflicting by specifying what matters and when it matters. The inconsistency can be explained by differences in the strength of agglomeration forces across industries, countries or time periods, but also by methodological issues and the various indicators of MAR and Jacobs externalities used in the research.

Moreover, the suitability of the specialized or diversified regions for particular industries was assessed in a more detailed analysis of industrial sectors. Although not overwhelming, some evidence was found that in low tech sectors, Marshall externalities have stronger effects than Jacobs externalities. The situation in medium tech sectors yields similar results for both theories, but differs for the high tech sectors. The latter slightly favour diversified regions, while the effects of Marshall externalities are less pronounced. Diversification also appears to be a growth promoter in services. Furthermore, it was shown that the role of externalities varies according to the maturity the industry. Jacobs externalities predominate in the early stages of the industry life cycle, whereas Marshall externalities enter at a later point, and in the end specialization will in fact hinder economic growth.

There are quite important implications of this investigation for public policy. In general, this thesis suggests that in regions with mature, low tech industries, regional policy should emphasize the development of a narrow set of economic activities in the region, which will presumably lead to greater productivity. In high tech regions, on the other hand, policy should focus on the creation of a diverse set of economic activities, which should enhance economic development. However, given such contrasting opinions and conflicting conclusions, any regional development policy which selects, supports or discriminates certain industrial activities or technologies should be applied with caution until the issue is fully clarified.

The main focus of the thesis is however more specific – it concerns Canadian high technology clusters. It was shown that innovative activity in Canada is concentrated in several locations which roughly correspond to the larger metropolitan areas. 12 biotechnology and 8 nanotechnology clusters have been identified. In biotechnology, more than half of all Canadian inventors reside in three largest clusters - Toronto, Montreal and Vancouver, but in nanotechnology it is mainly the Toronto cluster which dominates the industrial sector since around one quarter of all Canadian inventors live there. The thesis has made a contribution by making a profile description for the Canadian biotechnology clusters in terms of patenting quality and quantity, the nature of biotechnology activities, the properties of assignees and their propensity to collaborate.

Around half of the biotechnology patents are assigned to firms. However, publicly-funded research is highly important for biotechnology in Canada. Universities are the most active institutions in biotechnology and the greatest producers of patents. The production of patents is however very different among Canadian universities and several renowned research universities that are highly active in biotechnology research own only an inferior number of the patents. The existence, quality and effectiveness of the technology transfer support available at these universities as well as the university intellectual property rules and policies were both suggested to be the cause. The contribution of the government laboratories to the biotechnology research and development is also substantial. Biotechnology cluster-based subnetworks are better developed and the research is better organised in clusters which host the five NRC institutes (Montreal, Vancouver, Ottawa, Saskatoon and Halifax although the latter is a much smaller cluster). In contrast, in nanotechnology, only two poles are present, Toronto and Edmonton, the latter still emerging.

An interesting issue uncovered by this thesis is that although there is a great innovation capability among Canadian researchers, a lot of the intellectual property actually leaves the country. This is especially evident in nanotechnology. Almost half of all the innovations authored or co-authored by Canadian nanotechnology inventors are assigned to the foreign subjects. Although Canadians do the research, the fruit of their labour is not appropriated within Canada.

Even though Canada has a quite high share of the industry-financed research in academic sector, only very little evidence of cooperation among biotechnology companies and academic institutions was found. In fact, the patent co-assignment data suggest that Canadian institutions do not collaborate much in general, no matter what is the type of the institution and whether the collaborative ties lie within or outside clusters. The most frequent typical partner of a Canadian biotechnology institution for the pursuit of joint research activities is another institution abroad (mainly in the US).

Much more collaboration was detected when instead of the institutional cooperation the cooperative relationships among the individual inventors were examined. Most of the collaborative activity of Canadian inventors take place within Canadian clusters, while the inter-cluster collaboration in Canada is much less common for both biotechnology and nanotechnology inventors. As in institutional collaborations, international ties also account for the highest proportion of all the collaborations outside the clusters and the most popular foreign collaboration partners for Canadian inventors also reside south of the border, in the USA. It was interesting to observe that if Canadian inventors do not find collaboration partners inside their own clusters, they in fact prefer to collaborate with foreigners rather than to carry out the joint research with fellow Canadian inventors. Around one third of the inventors in both databases are foreign residents and they are so entangled into the collaboration network of Canadians that it was noticed that their presence is in fact critical for the transmission of knowledge between Canadian inventors themselves. Foreigners are extremely important in connecting Canadian inventors from different clusters (or even those from the same cluster) together.

When the observed magnitude of foreign collaboration is taken into consideration, the importance of the geographical distance for the decision on the joint research project partners may be questioned. However, the thesis concludes that the distance does play an important role in selecting the research collaborators in both biotechnology and nanotechnology. An overwhelming preference of the Canadian inventors is towards local and relatively proximate partnerships. Nonetheless, if the suitable collaborators are not found within the distance of 600 km, the importance of the geographical factor significantly decreases, since in this case both biotechnology and nanotechnology inventors quite often opt for very distant or overseas cooperation partners.

The collaborative structures within biotechnology and nanotechnology networks are quite distinct. The biotechnology innovation network is larger, more developed and less fragmented than the nanotechnology one. The higher fragmentation of the nanotechnology network is explained by the greater disparity among the nanotechnology specializations compared to the more closely related biotechnology fields. The distances in biotechnology network are longer, but so is the maximal reach, which enables bringing fresh and non-redundant knowledge from distant locations.

The architecture of the network of Canadian biotechnology inventors was investigated within two different concepts: First, collaboration among inventors working in clusters (geographical proximity); second, cooperation among inventors who are directly or indirectly interconnected in network components (cognitive proximity). It was noticed that knowledge transmission is more efficient through the network components, but as was already revealed above, most of collaborations in Canadian biotechnology still greatly depend on the geographical circumstances and take place within clusters. Both geographical space (based on clusters) and the technological space (based on network) are thus at utmost importance for the knowledge creation and diffusion. The geographical and technological dimensions both nurture the growth of the cluster and promote innovation through a dynamic interaction of the actors localized in clusters who absorb external knowledge through the local and non-local networks. The geographical and technological collaboration spaces thus overlap to a certain extent, but they differ in their structures: Many inventors from the same cluster may also be part of the same network component. The bulk of smaller components are entirely contained within one cluster, larger components however usually encompass several clusters. The points of interaction (inventors well connected both inside and outside the clusters) between the two collaboration spaces were then examined further.

Gatekeepers are the inventors who bridge over the geographical and technological spaces and hence enable the nurturing of biotechnology clusters with fresh external knowledge. This thesis proposes indicators, which measure each inventor's importance as a procurer of external knowledge for the cluster (or for Canada) based on the share of innovative production to which he thereby contributes. Only around 10%-20% of all inventors in most clusters were identified as gatekeepers who are responsible for the inflow of external information to the cluster. The patenting productivity of these inventors (and their coincidence with star scientists) was another question studied in this thesis.

Star scientists are recognized as a key driving force behind the growth and innovation in biotechnology. In order to identify the most prolific inventors in Canadian biotechnology clusters, new measures were proposed. These take into consideration only the patent quantity (star inventors), both the patent quantity and quality (*QQ-star inventors*), and the number of forward citations in scientific articles (highly cited scientists). These criteria then enabled to distinguish and compare various prominent inventors with some interesting conclusions: Star inventors or highly prolific biotechnology inventors do not necessarily author or co-author patents of the highest value. Furthermore, not all the highly cited researchers and scientists superior in the biotechnology field produce patents or register them at the USPTO.

Finally, the gatekeeping role of these prominent inventors was examined. The great majority of the star inventors, majority of *QQ-star inventors* and almost half of all the highly cited scientists were also identified as gatekeepers responsible for the inflow of external information which highly contributes to the Canadian innovative potential.

The results of this research are of great importance for Canada as the thesis focuses on two of its most dynamic fields, biotechnology and nanotechnology, which provide a significant contribution to science advancement and innovation, thousands of jobs, as well as large exports. These technologies are studied in order to understand the factors that favour innovation within clusters in Canada. The Conference Board of Canada (Munn-Venn and Voyer, 2004) has made a significant recommendation for government to support the development of clusters by investing in the knowledge infrastructure, by developing skilled labour and by promoting networking and research. By characterizing Canadian high tech clusters and shedding light on the knowledge transmission processes that are carried out through innovation networks, this thesis has greatly contributed to the understanding of knowledge transfers that characterise high technology sectors in Canada.

9.2 Recommendations for future research

In the core of this thesis lies the construction of the biotechnology (nanotechnology) innovation networks from the patent data found in the USPTO database. The understanding of the topic could be much improved if the sources of input information are enriched. Several avenues how to proceed exist, some of which are already being explored:

The network of Canadian biotechnology (nanotechnology) scientists, who are the authors or co-authors of the scientific articles, could be constructed. This would enable to compare the structure of the networks of various innovators (i.e. inventors and scientists) and its impact on the innovative propensity of the firms in clusters. The two databases could also be merged in order to follow how the ideas from basic or applied science (evidenced by articles) transform into the innovative products (evidenced by patents).

The citations of the patents could also be extracted from the USPTO database and added to complete the picture. Innovation could then be viewed as a continuous process of older inventions stimulating and facilitating future inventions. The patent citation patterns over time would allow following the path of knowledge diffusion and knowledge obsolescence. The information on patent citations would also serve as an indicator of the economic value of inventive activity, which could be compared with the indicator used in this thesis – the number of patent claims.

The information about the affiliation for each inventor could be gathered in order to better understand from which environment they arise (academics, industrial or governmental inventors) and what influence this has on the overall patent production in the cluster. Some revealing insights about the collaboration between academia and industry could thereby be obtained. The separate innovation networks for academic and industrial inventors then could be created and their structural properties explored. Moreover, it would be interesting to gain the information on the background of the prominent individuals - gatekeepers and star or *QQ-star inventors*.

Finally, all the worldwide biotechnology (nanotechnology) patents could be included in the study. This would enable to see the networks in their entirety and to acquire a full picture of innovation production in Canadian biotechnology (nanotechnology). As was suggested in this thesis, the Canadian inventors would probably be much more intertwined if both all their connections to all other inventors in the world and all the connections among these international inventors are considered at the same time.

REFERENCES

ACS, Z.J. AND AUDRETSCH, D.B. (1987). Innovation, market structure and firm size. Review of Economics and Statistics, 69, 567-575.

ACS, Z.J. AND AUDRETSCH, D.B. (1988). Innovation in large and small firms: An empirical Analysis. The American Economic Review, 78(4), 678-690.

ACS, Z.J. AND AUDRETSCH, D.B. (1990). Innovation and Small Firms. Cambridge, MA: MIT Press.

ACS, Z.J., ANSELIN, L., AND VARGA, A. (2002). Patent and innovation counts as measures of regional production of new knowledge. Research Policy, 31(1069-1085).

ACS, Z.J., AUDRETSCH, D.B., AND FELDMAN, M.P. (1992). Real effects of academic research: Comment. The American Economic Review, 82(1), 363-367.

ACS, Z.J., AUDRETSCH, D.B., AND FELDMAN, M.P. (1994). R&D spillovers and recipient firm size. The Review of Economics and Statistics, 76(2), 336-340.

AHARONSON, B.S., BAUM, J.A.C. AND FELDMAN, M.P. (2004). Industrial Clustering and the Returns to Inventive Activity: Canadian Biotechnology Firms, 1991-2000, Working paper, University of Toronto

ALBERT, R., JEONG, H., AND BARABASI, A.-L. (1999). Diameter of the world-wide web. Nature, 401, 130-131.

ALLEN, R.C. (1983). Collective invention. Journal of Economic Behaviour and Organization, 4, 1-24.

ALLEN, T.J. (1967). Communication in the research and development laboratory. Technology Review, 70 (1), 31-7.

ALLEN, T.J. (1977). Managing the Flow of Technology. Technology Transfer and the Dissemination of Technological Information within the R&D Organization. MIT Press, Cambridge.

ALMEIDA, P. AND KOGUT, B. (1997). The exploration of technological divrsity and the geographic localisation of innovation. Small Business Economics, 9, 21-31.

ALMEIDA, P. AND KOGUT, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. Management Science, 45, 905-916.

ALMEIDA, R. (2003). Local characteristics and growth in Portuguese regions. Working Paper (Universidad de Barcelona).

ALMEIDA, R. (2006). Local Economic Structure and Growth, World Bank Research Department, Working paper.

AMARAL, L., SCALA, A., BARTHELEMY, M., AND STANLEY, H. (2000). Classes of small-world networks. Proceedings of the National Academy of Sciences, 97(21), 11149-11152.

ANCORI, B., BURETH, A., AND COHENDET, P. (2000). The economics of knowledge: The debate about codification and tacit knowledge. Industrial and Corporate Change, 9(2), 255-287.

ANDERSSON, M. AND EJERMO, O. (2003). Knowledge production function in Swedish functional regions 1993-1999. CESPRI Working Papers, 139.

ANGEL, D.P. (2002). Inter-firm collaboration and technology development partnerships within US manufacturing industries. Regional Studies, 36(4), 333-344.

ANSELIN, L., VARGA, A., AND ACS, Z. (1997). Local geographic spillovers between university research and high technology innovations. Journal of Urban Economics, 42, 422-448.

ARCHER, N. AND WANG, S. (2002) Knowledge management in the network organization. Proceedings of the 24th International Conference on Information Technology Transfers (IEEE). 1, 299-304.

ARROW, K.J. (1962). The economic implications of learning by doing. The Review of Economic Studies, 29, 155-172.

ARTHUR, W.B. (1990). Silicon Valley locational clusters: Do increasing returns imply monopoly. Mathematical Social Sciences, 19, 235-251.

AUDRETSCH, D.B. (1995). Innovation and Industry Evolution. Cambridge, MA: MIT Press.

AUDRETSCH, D.B. (1998). Agglomeration and the location of innovative activity. Oxford Review of economic Policy, 14(2), 18-29.

AUDRETSCH, D.B. AND FELDMAN, M.P. (1996). R&D spillovers and the geography of innovation and production. The American Economic Review, 86(3), 630-640.

AUDRETSCH, D.B. AND LEHMANN, E.E. (2005). Does the knowledge spillover theory of entrepreneurship hold for regions? Research Policy, 34, 1191-1202.

AUDRETSCH, D.B. AND STEPHAN, P.E. (1996). Company-scientist locational links: The case of biotechnology. The American Economic Review, 86(3), 641-652.

AUDRETSCH, D.B., LEHMANN, E.E., and WARNING, S. (2005). University spillovers and new firm location. Research Policy, 34, 1113-1122.

AUTANT-BERNARD, C. (2001). The geography of knowledge spillovers and technological proximity, Economics of Innovation and New Technology, 10, 237-254.

BACHMANN, G. (1998). Innovationsschub aus dem Nanokosmos, VDI, Düsseldorf

BAILEY, R. (2003). The Smaller the Better, Reason, (Dec. 1), p. 44. Available on line: (http://www.reason.com/news/show/28969.html)

BALCONI, M., BRESCHI, S., AND LISSONI, F. (2004). Networks of inventors and the role of academia: An exploration of Italian patent data. Research Policy, 33, 127-145.

BAPTISTA, R. AND SWANN, P. (1998). Do firms in clusters innovate more? Research Policy, 27(525-540).

BAPTISTA, R. AND SWANN, P. (1999). A comparison of clustering dynamics in the US and UK computer industries. Journal of Evolutionary Economics, 9, 373-399.

BATEN, J., SPADAVECCHIA, A., YIN, S., AND STREB, J. (2005). Clusters, externalities and innovation: New evidence from German firms 1878-1913. Working paper, 2005/4.

BATISSE, C. (2002). Structure industrielle et croissance locale en république populaire de Chine, Revue Région et Développement, 16, 85-110.

BAUM, J.A., SHIPILOV, A.V., AND ROWLEY, T.J. (2003). Where do small worlds come from? Industrial and Corporate Change, 12, 697-725.

BAWA, R. (2004). Nanotechnology patenting in the US, Nanotechnology Law and Business, 1, 31–50.

BEARDSELL, M., AND HENDERSON, V. (1999). Spatial evolution of the computer industry in the USA, European Economic Review, 43, 431-456.

BEAUCAGE, J.-S. AND BEAUDRY, C. (2006). The Importance of Knowledge Networks within Canadian Biotechnology Clusters. International Schumpeter Conference, Sophia-Antipolis.

BEAUDRY, C. (2001). Entry, growth and patenting in industrial clusters. International Journal of Economics of Business, 8(3), 405-436.

BEAUDRY, C. and BRESCHI, S. (2000). Does 'clustering' really help firms' innovative activities? CESPRI, Working paper 111.

BEAUDRY, C. and BRESCHI, S. (2003). Are firms in clusters really more innovative. Economics of Innovation and New Technology, 12(4), 325-342.

BEAUDRY, C. and SWANN, G. (2008 forthcoming) Firm-level growth in industrial clusters: a bird's eye view of the United Kingdom, Small Business Economics.

BEAUDRY, C. and SWANN, P. (2001). Growth in industrial clusters: a bird's eye view of the United Kingdom, The Stanford Institute for Economic Policy Research at Stanford University, SIEPR Discussion Paper No. 00-38

BEAUDRY, C., BRESCHI, S. and SWANN, G. (2001). Clusters, innovation and growth: A comparative study of European countries, in: Dunning, J., Mucchielli, J.L. (Eds.), Multinational Firms: The Global and Local Dilemma, Routledge, London, UK, pp. 190-213.

BEAUMONT, J. (1982). The location, mobility and finance of new high-technology companies in the UK electronics industry. Department of Trade and Industry Report.

BERGER, M. (2006). The Patent Land Grab in Nanotechnology Continues Unabated, Creating Problems Down the Road, Nanowerk, (March 30). Available on line: (http://www.nanowerk.com/spotlight/spotid=386.php)

BHASKARABHATLA, A.S. (2006). Nanotechnology enterprise in the United States: structure and location, Thesis, Georgia Institute of Technology.

BLACK, D. AND HENDERSON, V. (1999). Spatial evolution of population and industry in the United States, AEA Papers and Proceedings 89, 321-327.

BONACCORSI, A. AND THOMA, G. (2006). Combinatorial Inventors and Performance in the Emerging Field of Nano Science and Technology, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Paris. Available on line: (http://www.nanodistrict.org/events/Nano%20session%20at%20the%20prime%20conference/nanotech_presentation%20prime%202nd%20annual%20conf.pdf)

BOSCHMA, R.A. (2005). Proximity and innovation: A critical assessment. Regional Studies, 39(1), 61-74.

BOSCHMA, R.A. AND TER WAL, A.L. J. (2007). Knowledge networks and innovative performance in an industrial district: The case of a footwear district in the South of Italy. Industry and Innovation, 14(2), 177.

BOSCHMA, R.A., NEFFKE, F.M.H., AND VAN OORT, F. (2005) Externalities and the industry life cycle. CEFR 5th Spring School in Economic Geography. 2005. University of Cagliari.

BRADLEY, R., AND GANS, J., 1998. Growth in Australian cities, Economic Record 74, 266-278.

BRAUN, T., SCHUBERT, A. AND ZSINDELY, S. (1997) Nanoscience and nanotechnology on the balance. Scientometrics, 38, 321-325.

BRESCHI, S. AND LISSONI, F. (2001a). Knowledge spillovers and local innovation systems: A critical survey. Industrial and Corporate Change, 10(4), 975-1005.

BRESCHI, S. AND LISSONI, F. (2001b). Localized knowledge spillovers vs. innovative milieux: Knowledge ``tacitness`` reconsidered. Papers in Regional Science, 80, 255-273.

BRESCHI, S. and LISSONI, F. (2003). Mobility and social networks: localized knowledge spillovers revisited. CESPRI Working Papers, 142.

BRESCHI, S. AND LISSONI, F. (2004). Knowledge networks from patent data: Methodological issues and research targets. CESPRI Working Papers, 150.

BRESCHI, S. AND MALERBA, F. (2001). The geography of innovation and economic clustering: Some introductory notes. Industrial and Corporate Change, 10(4), 817-833.

BURT, R. (2001). Bandwidth and echo: Trust, information, and gossip in social networks . In: A. Casella and J.E. Rauch (Eds.), Networks and Markets: Contributions from Economics and Sociology New York: Russel Sage Foundation.

BURT, R.S. (1980). Autonomy in a social topology. American Journal of Sociology. 85(4), 892-925

BURT, R.S. (1992). Structural Holes: The Social Structure of Competition. Harvard University Press: Cambridge, MA.

BURT, R.S. (2004). Structural holes and good ideas. American Journal of Sociology, 110(2), 349-399.

BURT, R.S., 2007. Secondhand brokerage: Evidence on the importance of local structure for managers, bankers, and analysts. Academy of Management Journal, 50(1), 119-148.

CABALLERO, R.J. AND JAFFE, A.B. (1993). How high are the giants` shoulders: An empirical assessment of knowledge spillovers and creative destruction in a model of economic growth. In: O. Blanchard and S. Fischer (Eds.), NBER Macroeconomics Annual MIT Press.

CAINELLI, G., LEONCINI, R., MONTINI, A., 2001. The evolution of industrial sectors in Europe, Nelson and Winter Conference, Aalborg, 12-15 June 2001.

CALLEJON, M AND COSTA, T. (1996) External economics and the location of an industry. 36th European Congress of European Regional Science Association. August 1996. Zurich, Switzerland.

CANTNER, U. AND GRAF, H. (2006). The network of innovators in Jena: An application of social network analysis. Research Policy, 35, 463-480.

CAPELLO, R. (2002). Spatial and sectoral characteristics of relational capital in innovation activity, European Planning Studies 10, 177-200.

CINGANO, F. AND SCHIVARDI, F. (2004). Identifying the sources of local productivity growth. Journal of the European Economic Association, 2, 720-742.

COCKBURN, I.M. AND HENDERSON, R.M. (1998). Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. The Journal of Industrial Economics, 46(2), 157-182.

COHEN, W.M. AND LEVINTHAL, D.A. (1989). Innovation and learning: The two faces of R&D. Economic Journal, 99, 569-596.

COHEN, W.M. AND LEVINTHAL, D.A. (1990). Absorptive capacity: A new perspective on learning and innovation. Administrative Science Quarterly, 35(1), 128-152.

COMBES, P.-P. (2000a). Economic Structure and Local Growth: France 1984-1993, Journal of Urban Economics, 47, 329-355.

COMBES, P.-P. (2000b). Marshall-Arrow-Romer externalities and city growth, CERAS Working paper, No. 99-06.

COMBES, P.-P., MAGNAC, T. AND ROBIN, J.-M. (2004). The dynamics of local employment in France, Journal of Urban Economics, 56, 217–243.

COSTA-CAMPI, M. AND VILADECANS-MARSAL, E. (1999). The District Effect and the Competitiveness of manufacturing Companies in Local Productive Systems, Urban Studies, 36, 2085-2098.

COTA, J. (2002). Agglomeration economies and urban manufacturing growth in the northern border cities of Mexico, El Colegio de la Frontera Norte Autopista Tijuana-Ensenada, Working paper.

COTA, J.E.M. (2002). Agglomeration economies and urban manufacturing growth in the northern border cities of Mexico. Working paper (El colegio de la frontera norte autopista Tijuana-Ensanada).

COWAN, R. AND FORAY, D. (1997). The economics of codification and the diffusion of knowledge. Industrial and Corporate Change, 6(3), 595-622.

COWAN, R. AND JONARD, N. (2003). The dynamics of collective invention. Journal of Economic Behaviour and Organization, 52, 513-532.

COWAN, R. AND JONARD, N. (2004). Network structure and the diffusion of knowledge. Journal of Economic Dynamics and Control, 28, 1557-1575.

COWAN, R., DAVID, P.A., AND FORAY, D. (2000). The explicit economics of knowledge codification and tacitness. Industrial and Corporate Change, 9(2), 211-253.

COWAN, R., JONARD, N. AND OZMAN, M. (2004). Knowledge dynamics in a network industry. Technological Forecasting and Social Change, 71, 469-484

COWAN, R., JONARD, N., AND ZIMMERMANN, J.-B. (2005). Bilateral collaboration and emergent networks. SSRN Working Papers.

COWAN, R., JONARD, N., AND ZIMMERMANN, J.-B. (2007). Bilateral collaboration and the emergence of innovation networks. Management Science, 53, 7, 1051-1067

CREA. (2005). Creative capabilities and the promotion of highly innovative research in Europe and the United States, European Commission, Project Reference 511889 (NEST).

CUMMINGS, T. (1991). Innovation through alliances: Archimedes' hall of mirrors. European Management Journal, 9, 284-287

DAHL, M.S. AND PEDERSEN, C.O.R. (2004). Knowledge flows through informal contacts in idustrial clusters: myth or reality? Research Policy, 33, 1673-1686.

DARBY, M.R. AND ZUCKER, L.G. (2003). Grilichesian breakthroughs: Inventions of methods of inventing in nanotechnology and biotechnology. NBER Working Papers, 9825.

DARBY, M.R. AND ZUCKER, L.G. (2004), Formation and transformation of industries: Nanotechnology, chapter 2 in Nanoscience and Nanotechnology: Opportunities and Challenges in California, Sacramento, CA: California Council on Science and Technology. Available on line: (http://www.ccst.us/publications/2004/2004NanoNC.pdf)

DASGUPTA, P. AND DAVID, P.A. (1994). Toward a new economics of science. Research Policy, 23, 487-521.

DAVIS, G.F., YOO, M., AND BAKER, W.E. (2003). The small world of the American corporate elite, 1982-2001. Strategic Organization, 1, 301-326.

DE LUCIO, J.J., HERCE, J.A. AND GOICOELA, A. (1996). Externalities and industrial growth. Spain: 1978-1992. 36th European Congress of European Regional Science Association. Zurich, Switzerland.

DE LUCIO, J.J., HERCE, J.A. AND GOICOLEA, A. (2002). The effects of externalities on productivity growth in Spanish industry, Regional Science and Urban Economics 32, 241–258.

DE NOOY, W., MRVAR, A. AND BATAGELJ, A. (2005). Exploratory Social Network Analysis with Pajek. Cambridge: Cambridge University Press

DEIDDA, S., PACI, R. AND USAI, S. (2006). Spatial Externalities and Local Economic Growth, Crenos and University of Sassari, Contributi di Ricerca CRENoS 02/06.

DEKLE, R. (2002). Industrial Concentration and Regional Growth: Evidence from the Prefectures, Review of Economics and Statistics 84, 310–315.

DERWENT, T. (2003). Web-based nanotechnology patent searches available, Canadian Chemical News, 55 (March 3)

DODDS, P., MUHAMAD, R., AND WATTS, D. (2003). An experimental study of search in global social networks. Science, 301, 827-829.

DUCOR, P. (2000). Intellectual property: Coauthorship and coinventorship. Science, 289(5481), 873-875.

DUGUET, E. AND MACGARVIE, M. (2005). How well do patent citations measure flows of technology: Evidence from French innovation surveys. Economics of Innovation and New Technology, 14(5), 375-393.

DURANTON, G. AND PUGA, D. (2001). Nursery Cities, American Economic Review 91, 1454–1475.

DURANTON, G. AND PUGA, D. (2004). Micro-foundations of urban agglomeration economies, in: Henderson, J.V., Thisse, J.-F., Handbook of Regional and Urban Economics, North-Holland, pp. 2063-2117.

EJERMO, O. (2005). Technological diversity and Jacobs externality hypothesis revisited. Growth and Change, 36(2), 167.

EJERMO, O. AND KARLSSON, C. (2006). Interregional inventor networks as studied by patent coinventorships. Research Policy, 35, 412-430.

ERNST AND YOUNG (2002). Beyond Borders: The Canadian Biotechnology Report 2002, Toronto.

ETC (2005). Nanotech's "Second Nature" Patents. ETC Group. (August 26). Available on line: (http://www.ipfrontline.com/depts/article.asp?id=5456anddeptid=5)

FELDMAN, M. (1994). The geography of innovation, Kluwer Academic Publishers, Boston.

FELDMAN, M.P. AND AUDRETSCH, D.B. (1999). Innovation in cities: Science-based diversity, specialization and localized competition. European Economic Review, 43, 409-429.

FERNANDEZ, R. M. AND GOULD, R.V. (1994). A dilemma of state power: Brokerage and influence in the national health policy domain. American Journal of Sociology, 99, 1455-1491.

FISHER, M.M. AND VARGA, A. (2003). Spatial knowledge spillovers and university research: Evidence from Austria. The Annals of Regional Science, 37(2), 303-322.

FLEMING, L., KING, C.IV., AND JUDA, A. (2006). Small worlds and innovation. SSRN Working Papers.

FLEMING, L., MINGO, S. AND CHEN, D. (2007). Collaborative brokerage, generative creativity, and creative success. Administrative Science Quarterly, 52(3), 443.

FORNI, M. AND PABA, S. (2002). Spillovers and the growth of local industries, The Journal of Industrial Economics. 50, 151-171.

FOSFURI, A. AND RONDE, T. (2004). High-tech clusters, technology spillovers, and trade secret laws. International Journal of Industrial Organization, 22, 45-65.

FRENKEN, K., VAN OORT, F. G., AND VERBURG, T. (2005) Variety and regional economic growth in the Netherlands. Regional Studies Conference on Regional Growth Agenda. 28-31 May 2005, Aalborg.

FURUKAWA, R. AND GOTO, A. (2006). The role of corporate scientists in innovation, Research Policy, 35, 24-36.

GAMBARDELLA, A. AND GARCIA-FONDES, W. (1996). European research funding and regional technological capabilities. Working Paper.

GAO, T. (2004). Regional industrial growth: evidence from Chinese industries, Regional Science and Urban Economics 34, 101–124.

GAUVIN, S. (1995). Networks of innovators: Evidence from Canadian patents. Group Decision and Negotiation, 4, 411-428.

GEROSKI, P. (1995). Markets for technology: Knowledge, innovation and appropriability. In: P. Stoneman (Ed.), Handbook of the Economics of Innovation and Technological Change (pp. 90-131). Oxford: Blackwell.

GITTELMAN, M. (2006). Does Geography Matter for Science-Based Firms? Epistemic Communities and the Geography of Research and Patenting in Biotechnology. DRUID Summer Conference on Knowledge, Innovation and Competitiveness: Dynamics of Firms, Networks, Regions and Institutions, Copenhagen, Denmark.

GLAESER, E.L., KALLAL, H.D., SCHEINKMAN, J.A., AND SHLEIFER, A. (1992). Growth in cities. Journal of Political Economy, 100, 1126-1152.

GOULD, R.V. AND FERNANDEZ, R.M. (1989). Structures of mediation: A formal approach to brokerage in transaction networks. Sociological Methodology, 19,89-126.

GOYAL, S., LEIJ, M., AND MORAGA-GONZALEZ, J. (2004). Economics: An emerging small world? Working Paper (Tinbergen Institute), TI2004-001/1.

GRAF, H., 2008. Gatekeepers in regional networks of innovators. The 25th DRUID Celebration Conference on Entrepreneurship and Innovation - Organizations, Institutions, Systems and Regions, Copenhagen, CBS, Denmark.

GREUNZ, L. (2004). Industrial structure and innovation: evidence from European regions, Journal of Evolutionary Economics 14, 563-592.

GRILICHES, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. The Bell Journal of Economics, 10, 93-116.

GRILICHES, Z. (1981). Market value, R&D and patents. Economic Letters, 7, 183-187.

GRILICHES, Z. (1984). R&D, Patents, and Productivity. Chicago: University of Chicago Press.

GRILICHES, Z. (1990). Patent statistics as economic indicators: A survey. Journal of Economic Literature, 28(4), 1661-1707.

GRILICHES, Z. (1992). The Search for R&D Spillovers, Scandanavian Journal of Economics 94, 29-47.

GUARE, J. (1990). Six Degrees of Separation: A play. New York: Vintage.

GUIMERA, R., UZZI, B., SPIRO, J., AND AMARAL, L.A.N. (2004). The assembly of creative teams and the emergence of the 'invisible college'. Working paper.

GULATI, R. AND GARGIULO, M. (1999). Where do interorganizational networks come from?. American Journal of Sociology, 104(1), 439-493.

HALL, B.H., JAFFE, A.B., AND TRAJTENBERG, M. (1998). Patent citations and market value: A first look. NBER Research Meeting.

HALL, B.H., JAFFE, A.B., AND TRAJTENBERG, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools. Working paper.

HANSON, G., (1998). Regional adjustment to trade liberalization, Regional Science and Urban Economics 28, 419–444.

HARADA, T., (2003). Three steps in knowledge communication: The emergence of knowledge transformers. Research Policy, 32: 1737–1751.

HARGADON, A. AND SUTTON R.I. (1997). Technology brokering and innovation in a product development firm. Administrative Science Quarterly, 42, 716-749.

HARHOFF, D., NARIN, F., SCHERER, F.M., AND VOPEL, K. (1999). Citation frequency and the value of patented inventions. The Review of Economics and Statistics, 81(3), 511-515.

HARRIS, R.G. (2001). The knowledge-based economy: intellectual origins and new economic perspectives. International Journal of Management Review, 3, 21-41.

HARRISON, B., KELLEY, M. AND GANT, J. (1996). Specialization versus diversity in local economies: The implications for innovative private-sector behaviour, Cityscape: A Journal of Policy Development and Research 2, 61-93.

HEINZE, T. (2004). Nanoscience and Nanotechnology in Europe: Analysis of Publications and Patent Applications including Comparisons with the United States, Nanotechnology Law and Business, 1(4), article 10. Available on line: (http://www.nanoandsociety.com/ourlibrary/documents/NanoInEurope.pdf)

HENDERSON, R. AND COCKBURN, I. (1994). Measuring competence? Exploring firm effects in pharmaceutical research. Strategic Management Journal, 15, 63-84.

HENDERSON, R. AND COCKBURN, I. (1996). Scale, scope and spillovers: The determinants of research productivity in drug discovery. RAND Journal of Economics, 27(1), 32-59.

HENDERSON, R., JAFFE, A., and Trajtenberg, M. (1998a). University patenting amid changing incentives for commercialization. In: G.B. Navaretti, P. Dasgupta, K.G. Maeler, and D. Siniscalco (Eds.), Creation and Transfer of Knowledge New York: Springer.

HENDERSON, R., JAFFE, A., AND TRAJTENBERG, M. (1998b). Universities as source of commercial technology: A detailed analysis of university patenting, 1965-1988. The Review of Economics and Statistics, 80(1), 119-127.

HENDERSON, V. (1986). Efficiency of resource usage and city size, Journal of Urban Economics 19, 47-70.

HENDERSON, V. (1997). Externalities and industrial development, Journal of Urban Economics 42, 449-470.

HENDERSON, V. (2001). Marshall's Scale Economies, Working paper CES-WP-01-17, December 2001.

HENDERSON, V., KUNCORO, A., AND TURNER, M. (1995). Industrial development in cities. The Journal of Political Economy, 103(5), 1067-1090.

HENDERSON, V., LEE, T. AND LEE Y. (2001). Scale Externalities in Korea, Journal of Urban Economics, 49, 479-504.

HICKS, D., BREITZMAN, D., OLIVASTRO, D., AND HAMILTON, K. (2001). The changing composition of innovative activities. Research Policy, 30(4), 681-703.

HOEN, A. (2001). Clusters: Determinants and effects. CPB Netherlands Bureau for Economic Policy Analysis Working Paper.

HOWELLS, J. (2006). Intermediation and the role of intermediaries in innovation. Research Policy, 35, 715–728.

HUANG *et al.* (2007). website of Artificial Intelligence Lab of University of California. Available on line: (http://128.196.40.92:8080/NanoMapper/Aboutdata.jsp)

HUANG, Z., CHEN, H., CHEN, Z.K. AND ROCO, M.C. (2004), International nanotechnology development in 2003: Country, institution and technology field analysis based on USPTO patent database, Journal of Nanoparticle Research, 6, 325–354.

HUANG, Z., CHEN, H., YIP, A., NG, G., GUO, F., CHEN, Z.K. AND ROCO, M.C. (2003), Longitudinal patent analysis for nanoscale science and engineering: Country, institution and technology field, Journal of Nanoparticle Research, 5, 333–363.

ISI (2002). ISI web site on nanotechnology: www.esi-topics.com/nano/

JACOBS, J. (1969). The economies of cities. NY : Random house.

JAFFE, A.B. (1986). Technological opportunity and spillovers of R&D: Evidence from firms patents, profits and market value. American Economic Review, 76(5), 984-1001.

JAFFE, A.B. (1989). Real effects of academic research. The American Economic Review, 79(5), 957-970.

JAFFE, A.B. (1998). Patents, patent citations, and the dynamics of technological change. NBER Reporter, 8-11.

JAFFE, A.B. AND LERNER, J. (2001). Reinventing public R&D: Patent policy and the commercialization of national laboratory technologies. RAND Journal of Economics, 32, 167-198.

JAFFE, A.B. AND TRAJTENBERG, M. (1996) Flows of knowledge from universities and federal laboratories. Modeling the flow of patent citations over time and across institutional and geographic boundaries. Proceedings of the National Academy of Sciences. 93.

JAFFE, A.B. AND TRAJTENBERG, M. (1999). International knowledge flows: Evidence from patent citations. Economics of Innovation and New Technology, 8, 105-136.

JAFFE, A.B. AND TRAJTENBERG, M. (2000). The meaning of patent citations: Report on the NBER/Case-Western Reserve Survey of Patentees. NBER Working Papers , 7631.

JAFFE, A.B., FOGARTY, M.S., AND BANKS, B.A. (1998). Evidence from patents and patent citations on the impact of NASA and other federal labs on commercial innovation. The Journal of Industrial Economics, 46(2), 183-205.

JAFFE, A.B., TRAJTENBERG, M., AND FOGARTY, M.S. (2000). Knowledge spillovers and patent citations: Evidence from a survey of inventors. The American Economic Review, 90(2), 215-218.

JAFFE, A.B., TRAJTENBERG, M., AND HENDERSON, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics, 108, 577-598.

JESPERSEN, S., SOKOLOV, I.M., AND BLUMEN, A. (2000). Small-world Rouse networks as models of cross-linked polymers. The Journal of Chemical Physics, 113(17), 7652-7655.

KAMEYAMA, Y. (2003). Dynamic externalities and the growth of manufacturing employment in Japanese cities: The roles of specialization and diversity. Working Paper (The International Centre for the Study of East Asian Develompment).

KANAMA, D. (2006). Patent application trends in the field of nanotechnology. Quarterly Review, 21, (October)

KATZ, R. AND TUSHMAN, M.L. (1981). An investigation into the managerial roles and career paths of gatekeepers and project supervisions in a major R&D facility. R&D Management, 11, 103-110.

KEEBLE, D. (1988). High-technology industry and local environments in the United Knogdom. In: P. Aydalot and D. Keeble (Eds.), High-technology industry and innovative environments: The European Experience London: Routledge.

KELLEY, M. AND HELPER, S. (1997). Firm size and capabilities, regional agglomeration, and the adoption of new technology, Carnegie Mellon University, Working paper.

KELLY, M. AND HAGEMAN, A. (1999). Marshallian externalities in innovation. Journal of Economic Growth, 4, 39-54.

KEPPLINGER, E.M. (2004). Patent Office Initiatives in nanotechnology, Presentation of American Intellectual Property Law Association. Available on line: http://www.aipla.org/Content/ContentGroups/Speaker_Papers/Spring_Meeting/20045/keppling_esther.pdf

KING III, C., SILK, A. AND KETELHÖHN, N. (2002). Knowledge spillovers and growth in the disagglomeration of the U.S. advertising agency industry, Harvard Business School Marketing Research Papers.

KIRAT, T. AND LUNG, Y. (1999). Innovation and proximity. Territories as loci of collective learning processes. European Urban and Regional Studies, 6(1), 27-38.

KOGUT, B. AND WALKER, G. (2001). The small world of Germany and the durability of national networks. American Sociological Review, 66, 317-335.

KORTE, C. AND MILGRAM, S. (1970). Acquaintance linking between white and negro populations: Application of the small world problem. Journal of Personality and Social Psychology, 15, 101-118.

KOSTOFF, R.N., STUMP, J.A., JOHNSON, D., MURDAY, J.S., LAU, C.G.Y. AND TOLLES,W.M. (2006). The structure and infrastructure of the global nanotechnology literature. Journal of Nanoparticle Research 8 (3–4), 301–321.

KREBS, V. (2006). Social network analysis. Available on line: (http://www.orgnet.com/sna.html)

KRUGMAN, P. (1991). Geography and Trade, Cambridge, MA: The MIT Press.

LAMOREAUX, N.R. AND SOKOLOFF, K.I. (1997). Location and technological change in the American glass industry during the late nineteenth and early twentieth century. NBER Working Papers, 5938.

LANJOUW, J.O. AND SCHANKERMAN, M. (2004). Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators. The Economic Journal, 114 (April), 441–465.

LAZARIC, N., LONGHI, C. AND THOMAS, C. (2008). Gatekeepers of knowledge versus platforms of knowledge: From potential to realized absorptive capacity. Regional Studies, 42(6), 837.

LE BLANC, G. (2000) Regional specialization, local externalities and clustering in information technology industries. XII Villa Mondragone International Economic Seminar 'Knowledge Economy, information Technologies and Growth'. 2000. University Tor Vergata, Roma.

LEE, B.; SOSIN, K. AND HONG, S. (2005). Sectoral manufacturing Productivity Growth in Korean Regions, Urban Studies, 42, 1201–1219.

LEE, L. L., CHAN, C. K., NGAIM, M. AND RAMAKRISHNA, S. (2006). Nanotechnology patent landscape, Nano, 1 (2), 101–113.

LI, X., LIN, Y., CHEN, H AND ROCO, M.C. (2007). Worldwide nanotechnology development: a comparative study of USPTO, EPO, and JPO patents (1976–2004), Journal of Nanoparticle Research, 9, 977–1002.

LIEBESKIND, J., OLIVER, A., ZUCKER, L., AND BREWER, M. (1996). Social networks, learning, and flexibility: Sourcing scientific knowledge in new biotechnology firms. Organization Science, 3, 783-831.

LINK, A.N. AND REES, J. (1990). Firm size, university based research and the returns to R&D. Small Business Economics, 2, 25-32.

LINK, A.N. AND SCOTT, J.T. (2005). Opening the ivory tower`s door: An analysis of the determinants of the formation of U.S. university spin-off companies. Research Policy, 34, 1106-1112.

LISSONI, F. (2001). Knowledge codification and the geography of innovation: The case of Brescia mechanical cluster. Research Policy, 30, 1479-1500.

LOIKKANEN, H. AND SUSILUOTO, I. (2002). An evaluation of economic efficiency of Finnish regions by DEA and Tobit models, 42st Congress of the European Regional Science Association, Dortmund, Germany, 27–31 August 2002.

LUX RESEARCH (2005). Nanotechnology gold rush yields crowded, entangled patents, Report from Lux Research and Foley and Lardner LLP. Available on line:
(http://www.luxresearchinc.com/press/RELEASE_IPreport.pdf)

MALIPIERO, A., MUNARI, F. AND SOBRERO, M. (2005). Focal firms as technological gatekeepers within industrial districts: Knowledge creation and dissemination in the Italian packaging machinery industry. DRUID Working Paper No. 05-05

MANO Y. AND OTSUKA, K. (2000). Agglomeration Economies and Geographical Concentration of Industries: A Case Study of manufacturing Sectors in Postwar Japan, Journal of the Japanese and International Economies, 14, 189–203.

MANSFIELD, E. (1984). Comment on using linked patent and R&D data to measure interindustry technology flows. In: Z. Griliches (Ed.), R&D, Patents, and Productivity (pp. 462-464). Chicago: University of Chicago Press.

MANSFIELD, E. (1991). Sources and characteristics of academic research underlying industrial innovations. Working Paper (University of Pensylvania).

MANSFIELD, E. (1995). Academic research underlying industrial innovations: Sources, characteristics and financing. The Review of Economics and Statistics, 77(1), 55-65.

MANSFIELD, E. (1998). Academic research and industrial innovation: An update of empirical findings. Research Policy, 26, 773-776.

MARIANI, M. (2000). Networks of inventors in the chemical industry. MERIT Research Memorandum.

MARINOVA, D. and McAleer, M. (2003) Nanotechnology strength indicators: International rankings based on US patents. Nanotechnology, 14, 1-7

MARSDEN, P.V. (1982). Brokerage behavior in restricted exchange networks. In: Marsden, P. and Lin, N. (Eds.), Social Structure and Network Analysis, Beverly Hills: Sage.

MARSHALL, A. (1890). Principles of Economics, London: MacMillan.

MARTIN, R. AND SUNLEY, P. (2003). Deconstructing clusters: Chaotic concept or policy panacea. Journal of Economic Geography, 3, 5-35.

MASKELL, P., ESKELINEN, H., HANNIBALSSON, I., MALMBERG, A., AND VANTE, E. (1998). Competitiveness, Localised Learning and Regional Development. London: Routledge.

MASSARD, N. AND RIOU, S. (2002). L'impact des structures locales sur l'innovation en France: Spécialisation ou diversité? Régions et développement, 16, 111-136.

MAURSETH, P.B. AND VERSPAGEN, B. (1999a). Knowledge Spillovers in Europe and its Consequences for Systems of Innovations. Aldershot: Edward Elgar publishing.

MAURSETH, P.B. AND VERSPAGEN, B. (1999b) Knowledge spillovers in Europe. A patent citation analysis. CRENOS Conference on Technological Externalities and Spatial Location. September 24-25 1999. University of Cagliari.

MCMILLAN, S.G., NARIN, F., AND DEEDS, D.L. (2000). An analysis of the critical role of public science in innovation: the case of biotechnology. Research Policy, 29, 1-8.

MEDOFF, M.H. (1996). A citation-based analysis of economists and economics programs. American Economist, 40(1), 46-59.

MEYER, M. (2001), Patent citation analysis in a novel field of technology: An exploration of nano-science and nano-technology. Scientometrics, 51, 163–183.

MILGRAM, S. (1967). The small world problem. Psychology Today, 2, 60-67.

MOGOUTOV, A. AND KAHANE, B., (2007). Data search strategy for science and technology emergence: a scalable and evolutionary query for nanotechnology tracking. Research Policy, 36, 893–903.

MONSENY, J.J. (2005). On the scope of agglomeration economies: Evidence from Catalan zip codes. Working Paper (Institut d'Economia de Barcelona).

MORONE, P. AND TAYLOR, R. (2004). Knowledge diffusion dynamics and network properties of face-to-face interactions. Journal of Evolutionary Economics, 14, 327-351.

MORRISON, A. (2008). Gatekeepers of knowledge within industrial districts: Who they are, how they interact. Regional Studies, 42(6), 817.

MUKKALA, K. (2004). Agglomeration economies in the Finnish manufacturing sector. Applied Economics, 36(21), 2419-24127.

MUNN-VENN, T. AND MITCHELL, P. (2005). Biotechnology in Canada: A Technology Platform for Growth, the report from the Conference Board of Canada. Available on line: (http://www.agwest.sk.ca/biotech/documents/115-06-Biotechnology%20in%20Canada.pdf)

MURRAY, F. (2002). Innovation as co-evolution of scientific and technological networks: Exploring tissue engineering. Research Policy, 31, 1389-1403.

NAKAMURA, R. (1985). Agglomeration economies in urban manufacturing industries: A case of Japanese cities, Journal of Urban Economics, 17, 108-124.

NARIN, F., HAMILTON, K., AND OLIVASTRO, D. (1997). The increasing linkage between US technology and public science. Research Policy, 26, 317-330.

NATIONAL RESEARCH COUNCIL CANADA (2005). Scan of Canadian Strengths in Biotechnology. Available on line: (http://www.science-metrix.com/pdf/SM_2004_013_NRC_Biotechnology_ Canadian_Strengths.pdf)

NCI (2006). Where science and law meet. Nanotechnology and intellectual property issues, National Cancer Institute, NCI Alliance for Nanotechnology in Cancer, (October). Available on line: (http://nano.cancer.gov/news_center/monthly_feature_2006_oct.pdf)

NEWMAN, M.E.J. (2000). Models of the small world. Journal of Statistical Physics, 101(3/4), 819-841.

NEWMAN, M.E.J. (2001a). Scientific collaboration networks. I. Network construction and fundamental results. Physical Review, 64, 016131.

NEWMAN, M.E.J. (2001b). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. Physical Review, 64, 016131.

NEWMAN, M.E.J. (2001c). Clustering and preferential attachment in growing networks. Physical Review, 64(025102).

NEWMAN, M.E.J. (2001d). The structure of scientific collaboration networks. Proceedings of National Academy of Sciences, 98(2), 404-409.

NEWMAN, M.E.J., WATTS, D.J., AND STROGATZ, S.H. (2002). Random graph models of social networks. Proceedings of the National Academy of Sciences, 99, 2566-2572.

NIOSI J. AND BANIK M. (2005). The evolution and performance of biotechnology regional systems of innovation. Cambridge Journal of Economics, 29, 343–357

NIOSI J. AND BAS, T.G. (2001). The Competencies of Regions – Canada's Clusters in Biotechnology. Small Business Economics, 17(1-2), 31-42.

NIOSI, J. (2003). Alliances are not enough explaining rapid growth in biotechnology firms. Research Policy, 32, 737-750.

NIOSI, J. (2005). Canada's Regional Innovation Systems. The Science-based Industries. McGill-Queen's University Press

NIOSI, J. (2006). Success factors in Canadian academic spin-offs. Journal of Technology Transfer, 31, 451–457.

NOYONS, C.M., BUTER, R.K., VAN RAAN, A.F.J., SCHMOCH, U., HINZE, S. AND RANGNOW, R. (2003). Mapping excellence in Science and Technology accross Europe in Nanoscience and nanotechnologies. Draft report of project EC-PPN CT-2002-0001 to the European Commission, Brussels. Available on line: (http://www.nanodistrict.org/bibliography/rapports/Final_report_13112003_nano.pdf)

OECD (2005). A Framework for Biotechnology Statistics. Available on line: (http://www.oecd.org/dataoecd/5/48/34935605.pdf)

OECD (2007). Main Science and Technology Indicators, Volume 2007/2

OLIVER, A.L. (1994). In between markets and hierarchies – networking through the life cycle of new biotechnology firms, ISSR Working Papers in the Social Sciences

ORSENIGO, L., PAMMOLI, F., AND RICCABONI, M. (2001). Technological change and network dynamics. Lessons from the pharmaceutical industry. Research Policy, 30, 485-508.

OUWERSLOOT, H. AND RIETVELD, P. (2000). The geography of R&D: tobit analysis and a Bayesian approach to mapping R&D activities in teh Netherlands. Environment and Planning A, 32, 1673-1688.

OWEN-SMITH, J. AND POWELL, W.W. (2004). Knowledge networks as channels and conduits: The effects of spillovers in the Boston biotechnology community. Organization Science, 15(1), 5-21.

PACI, R. AND USAI, S. (2000). Externalities, knowledge spillovers and the spatial distribution of innovation. GeoJournal, 49, 381-390.

PACI, R. AND USAI, S. (2005). Agglomeration economies and growth in Italian local labour systems. 1991-2001. Working paper.

PACI, R. AND USAI, S. (1999). The role of specialisation and diversity externalities in the agglomeration of innovative activities, University of Cagliari and CRENoS, Working paper.

PAKES, A. (1985). On patents, R&D, and the stock market rate of return. Journal of Political Economy, 93(2), 390-409.

PORTER, A., YOUTIE, J. AND SHAPIRA, P. (2006). Refining search terms for nanotechnology, Center for Nanotechnology in Society, Georgia Institute of Technology

PORTER, M. (2003). The economic performance of regions, Regional Studies, 37, 549-578.

PORTER, M.E. (1998). Clusters and the new economics of competition. Harvard Business Review, 77-90.

PREVEZER, M. (1997). The dynamics of industrial clustering in biotechnology. Small Business Economics, 8, 155-271.

PUTSCH, F. (2006). Analysis and modeling of science collaboration networks. Working Paper.

QUEENTON, J. AND NIOSI, J. (2003). Bioscientists and biotechnology: A Canadian study, 3rd European Meeting on Applied evolutionary Economics, Augsburg, Germany

RASMUSSEN, B. (2004). An Analysis of The Biomedical Sectors in Australia and Canada in a National Innovation Systems Context, Centre for Strategic Economic Studies, Victoria University of Technology, Working Paper Series, No. 21. Available on line:(http://www.cfses.com/documents/pharma/21-Biomedical_Sect_Aust_and_Can_Innovation_Rasmussen.pdf)

READ, C. (2005). Survey of intellectual property commercialization in the higher education sector in 2003, Statistics Canada

RIGBY, D.L. AND ESSLETZBICHLER, J. (2002). Agglomeration economies and productivity differences in US cities, Journal of Economic Geography 2, 407-432.

ROMER, P.M. (1986). Increasing returns and long-run growth. Journal of Political Economy, 94, 1001-1037.

RONDE, T. (2001). Trade secrets and information sharing. Journal of economics and Management Strategy, 10, 391-417.

ROSE, A. (2000). A Challenge for Measuring Biotechnology Activities, The Economics and Social Dynamics of Biotechnology, 2000.

ROSENTHAL, S. AND STRANGE, W. (2003). Geography, industrial organization, and agglomeration, The Review of Economics and Statistics, 85, 377–393.

ROTHAERMEL, F.T. AND THURSBY, M. (2007). The nanotech vs. the biotech revolution, sources of productivity in incumbent firm research. Research Policy, 36, 832–849.

SAMPAT, B. N. (2004) Examining Patent Examination: An Analysis of Examiner and Applicant Generated Prior Art, Haas Business School Working Paper. Available on line: (http://faculty.haas.berkeley.edu/wakeman/ba297tspring05/Sampat.pdf)

SAXENIAN, A. (1994). Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Cambridge, MA: Harvard University Press.

SCALA, A., AMARAL, L.A.N., AND BARTHÉLÉMY, M. (2001). Small-world networks and the conformation space of a short lattice polymer chain. Europhysics Letters, 55(4), 594-600.

SCHARTINGER, D., SCHIBANY, A., AND GASSLER, H. (2001). Interactive relations between universities and firms: Empirical evidence for Austria. Journal of Technology Transfer, 26(3), 255-268.

SCHERER, F.M. (1983). The propensity to patent. International. Journal of Industrial Organization, 1, 107-128.

SCHERER, F.M. (1991). Changing perspectives on the firm size problem. In: Z.J. Acs and D.B. Audretsch (Eds.), Innovation and Technological Change, (pp. 24-38). Ann Arbor: University of Michigan Press.

SCHIFFAUEROVA, A. AND BEAUDRY, C. (2008). Geographical Aspects of Collaboration in Canadian Nanotechnology Innovation, Nanobank Research Conference, Boston, MA, May 3, 2008.

SCHILLING, M.A. AND PHELPS, C.C. (2005). Interfirm collaboration networks: The impact of small world connectivity on firm innovation. SSRN Working Papers.

SCHILLING, M.A. AND PHELPS, C.C. (2007). The impact of large-scale network structure on firm innovation. Management Science, 53, 7, 1113-1126

SCHRADER, S. (1991). Informal technology transfer between firms: Cooperation through information trading. Research Policy, 20, 153-170.

SHEFER, D. AND FRENKEL, A. (1998). Local milieu and innovations: Some empirical results, The Annals of Regional Science, 32, 185-200.

SINGH, J. (2004). Innovation and knowledge diffusion in the global economy. Dissertation. Harvard School of Business Administration and Department of Economics.

SINGH, J. (2005). Collaborative networks as determinants of knowledge diffusion. Management Science, 51(5), 756-770.

SORENSON, O. (2003). Social networks and industrial geography. Journal of Evolutionary Economics, 13(5), 513-527.

SORENSON, O. AND STUART, T.E. (2001). Syndication networks and the spatial distribution of venture capital investments. American Journal of Sociology, 106(6), 1546-88.

STATISTICS CANADA (2001). Practices and Activities of Canadian Biotechnology Firms: Results from the Biotechnology Use and Development Survey - 1999. Available on line: (http://www.statcan.ca/english/research/88F0006XIE/88F0006XIB2001011.pdf)

STATISTICS CANADA (2002). Survey of Innovation 1999. Statistical Tables. Manufacturing industries. Available on line: (http://www.statcan.ca/english/research/88F0006XIE/88F0006XIE2002016.pdf)

STATISTICS CANADA (2005). Overview of the Biotechnology Use and Development Survey -2003. Available on line: (http://www.statcan.ca/english/research/ 88F0006XIE/88F0006XIE2005009.pdf)

STATISTICS CANADA (2007). Preliminary Results of the Biotechnology Use and Development Survey -2005. Available on line: (http://www.statcan.ca/Daily/English/070130/d070130c.htm)

STEINER, M. AND PLODER, M. (2008). Structure and strategy within heterogeneity: Multiple dimensions of regional networking. Regional Studies, 42(6), 793.

STIGLER, G.J. AND FRIEDLAND, C. (1975). The citation practices of doctorates in economics. Journal of Political Economy, 83, 477-507.

STRACHAN, G. (1995). The impact of regulations on the growth and evolution of Canadian commercial biotechnology. Current Opinion in Biotechnology, 1995, 6:261-263

STUART, T.E., OZDEMIR, S.Z. AND DING, W.W. (2007). Vertical alliance networks: The case of university–biotechnology–pharmaceutical alliance chains. Research Policy, 36, 477–498.

SUEDEKUM, J. AND BLIEN, U. (2005). Local Economic Structure and Industry Development in Germany, 1993–2001, Economics Bulletin, 15, 1–8.

SWANN, P. AND PREVEZER, M. (1996). A comparison of the dynamics of industrial clustering in computing and biotechnology. Research policy, 25, 1139-1157.

SWANN, P., PREVEZER, M. AND STOUT, D. (Eds.) (1998). The Dynamics of Industrial Clusters: International Comparisons in Computing and Biotechnology, Oxford University Press, Oxford.

THOMPSON, P. (2003). Patent citations and the geography of knowledge spillovers. Working paper (Carnegie university).

TOLLES, W.(2003). National security aspects of nanotechnology, In: M. C. Roco, W. S. Bainbridge (Eds), Societal Implications of Nanotechnology. NSET Workshop Report, NSF, Arlington, VA, pp. 173.187.

TONG, X. AND FRAME, J.D. (1994). Measuring national technological performance with patent claims data. Research Policy, 23,133-141

TORRE, A. AND GILLY, J-P. (2000). On the analytical dimension of proximity dynamics. Regional Studies, 34(2), 169-180.

TRAJTENBERG, M. (1990). A penny for your quotes: Patent citations and the value of innovations. The RAND Journal of Economics, 21(1), 172-187.

TRAJTENBERG, M., HENDERSON, R., AND JAFFE, A.B. (1992). Ivory tower versus corporate lab: An empirical study of basicness and appropriability. NBER Working Papers, 4146.

TRAJTENBERG, M., HENDERSON, R., AND JAFFE, A.B. (1997). University versus Corporate Patents: A Window on the Basicness of Invention. Economics of Innovation and New Technology, 5, 19-50.

TUSHMAN, M.L. AND KATZ, R. (1980). External communication and project performance: An investigation into the role of gatekeepers. Management Science, 26, 1071-1085.

USAI, S. AND PACI, R. (2003). Externalities and local economic growth in manufacturing industries. In: B. Fingleton (Ed.), European Regional Growth Berlin: Springer-Verlag.

UZZI, B. AND SPIRO, J. (2005). Collaboration and creativity: The small world problem. American Journal of Sociology, 111(2), 447-504.

VAN DER PANNE, G. (2004). Agglomeration externalities: Marshall versus Jacobs. Journal of Evolutionary Economics, 14(5), 593-604.

VAN DER PANNE, G. and VAN BEERS, C. (2006). On the Marshall – Jacobs controversy: it takes two to tango, Industrial and Corporate Change,15, 877-890.

VAN OORT, F.G. (2002). Innovation and agglomeration economies in the Netherlands, Journal of Economic and Social Geography, 93, 344–360.

VAN OORT, F.G. AND STAM, E. (2006). Agglomeration Economies and Entrepreneurship in the ICT Industry, Erasmus Research Institute of Management, Report Series ERS-2006-016-ORG.

VAN SOEST, P., GERKING, S.D., AND VAN OORT, F.G. (2002). Knowledge externalities, agglomeration economies, and employment growth in Dutch cities. CENTER Discussion paper, 2002-41.

VAN ZEEBROECK, N. AND VAN POTTELSBERGHE DE LA POTTERIE, B. (2006). Filing strategies and patent value. Solvay Business School, Working papers CEB

VARGA, A. (2000). Local academic knowledge transfers and the concentration of economic activity. Journal of Regional Science 40, 289-309.

VERSPAGEN, B AND SCHOENMAKERS, W. (2000) The spatial dimension of knowledge spillovers in Europe: Evidence from patenting data. AEA Conference on Intellectual Property Econometrics. April 19-20, 2000. Alicante.

VERSPAGEN, B. AND DUYSTERS, G. (2004). The small worlds of strategic technology alliances. Technovation, 24, 563-571.

VILADECANS-MARSAL, E. (2000). External economies and location of industrial activities: An analysis of the Spanish case, The 40th European Regional Science Association Congress, Barcelona.

VILADECANS-MARSAL, E. (2004). Agglomeration economies and industrial location: city-level evidence, Journal of Economic Geography, 4, 565-582.

VON HIPPEL, E. (1987). Cooperation between rivals: Informal know-how trading. Research Policy, 16, 291-302.

WAGNER, A. AND FELL, D.A. (2001). The small world inside large metabolic networks. Proceedings of the Royal Society B: Biological Sciences, 268(1478), 1803-1810.

WARRIS, C. (2004). Nanotechnology benchmarking project. Australian Academy of Science.

WASSERMAN, S. AND FAUST, K. (1994). Social network analysis. Cambridge: Cambridge University Press.

WATTS, D.J. (1999). Networks, dynamics, and the small-world phenomenon. American Journal of Sociology, 105, 4930527.

WATTS, D.J. AND STROGATZ, S.H. (1998). Collective dynamics of small-world networks. Nature, 393(4), 440-442.

WINCH, G. M. AND COURTNEY, R., 2007. The organization of innovation brokers: An international review. Technology Analysis and Strategic Management, 19(6), 747-763.

WINK, R. (2008). Gatekeepers and proximity in science-driven sectors in Europe and Asia: The case of human embryonic stem cell research. Regional Studies, 42(6), 777.

WOLFE, D.A. AND GERTLER, M.S. (2004). Clusters from the inside and out: Local dynamics and global linkages. Urban Studies, 41(5/6), 1071-1093.

WONG, P.K., HO, Y.P. AND CHAN, C.K. (2007). Internationalization and evolution of application areas of an emerging technology: The case of nanotechnology, Scientometrics, 70(3), 715–737

ZUCKER, L.G. AND BREWER, M.B. (1994). Intellectual capital and the birth of US biotechnology enterprises. NBER Working Papers, 4653.

ZUCKER, L.G. AND DARBY, M.R. (1996a). Costly information: Firm transformation, exit pr persistent failure. American Behavioral Scientist, 39, 959-974.

ZUCKER, L.G. AND DARBY, M.R. (1996b) Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. Proceedings of the National Academy of Sciences, 93, 12709-12716. 1996.

ZUCKER, L.G. AND DARBY, M.R. (2005). Socio-economic impact of nanoscale science: Initial results and nanobank. NBER Working Papers, 11181.

ZUCKER, L.G., DARBY, M.R. AND ARMSTRONG, J. (1998a). Geographically localized knowledge: Spillovers or markets? Economic Inquiry, 36(1), 65-86.

ZUCKER, L.G., DARBY, M.R. AND BREWER, M.B. (1996). Collaboration Structure and Innovation Dilemmas in Biotechnology. In: R.M.T.T.R. Kramer (Ed.), Trust in Organizations (pp. 90-113). Thousand Oaks, California: Sage.

ZUCKER, L.G., DARBY, M.R. AND BREWER, M.B. (1998b). Intellectual human capital and the birth of US biotechnology enterprises. American Economic Review, 88, 290-306.

ZUCKER, L.G., DARBY, M.R., FURNER, J., LIU, R.C. AND MAC, H. (2007). Minerva unbound: Knowledge stocks, knowledge flows and new knowledge production, Research Policy, 36, 850–863

APPENDICES

APPENDIX A (Chapter 2): EXAMPLE OF QUESTIONNAIRE ANSWERED BY A NANOTECHNOLOGY EXPERT

IS THIS A NANOTECHNOLOGY RELATED PATENT?

Each row in the table includes a title and an abstract of one patent. As an answer to the above question, please choose one of the following options:

- YES: The patent seems to be a nanotechnology related patent.
- PROBABLY YES: Nothing directly suggests that this is a nanotechnology relevant patent, but there is some nanotechnology work done in this topic.
- PROBABLY NO: Nothing suggests that this is a nanotechnology relevant patent and I am not sure whether nanotechnology includes this theme or not.
- NO: To my knowledge nanotechnology does not include this theme.
- I DON'T KNOW: It is absolutely impossible for me to evaluate the nanotechnology relevancy from the information provided.

Please, place a mark in one of the columns for each patent

	TITLE and ABSTRACT of the patent	YES	PROBABLY YES	PROBABLY NO	NO	I DON'T KNOW
1	**Method of expressing genes in mammalian cells** *Abstract* The present invention relates to a method of expressing a heterologous gene in mammalian cells and a recombinant DNA construct for use in the method. The invention also relates to a method of specifically killing cells which constitutively express AFP.				x	
2	**Monoclonal antibody to human cardiac myoglobin** *Abstract* A monoclonal antibody having high affinity to human cardiac myoglobin, which has undergone a conformational change resulting from the binding of the molecule to another molecule is described. This monoclonal antibody can be used in a rapid format double antibody immunoassay system to identify blood, serum or plasma levels of cardiac myoglobin. Such an immunoassay system can be used for diagnosing and quantifying myocardial necrosis and infarction.				x	

	TITLE and ABSTRACT of the patent	YES	PROBABLY YES	PROBABL Y NO	NO	I DON'T KNOW
3	**Lubricious coatings for substrates** *Abstract* The invention provides methods and kits to form water swellable gel coatings, preferably lubricious coatings, on substrates, and coated substrates thus formed. The coatings contain one or more antimicrobial metals formed with atomic disorder, together with one or more antimicrobial metals formed with atomic disorder such that the coatings provide an antimicrobial and anti-inflammatory effect when wet. The invention also provides a method to produce metal powders by sputtering a coating onto a moving surface, and then scraping the coating with one or more scrapers to produce the metal powder. The method is particularly useful for producing large amounts of nanocrystalline antimicrobial metal powders formed with atomic disorder, useful in the water swellable gel coatings of this invention.			X		
4	**Immunoassay diagnostic kit containing antigens derived from self-assembled, non-infectious, non-replicating, immunogenic retrovirus-like particles comprising modified HIV genomes and chimeric envelope glycoproteins** *Abstract* This invention is directed toward diagnostic kits and antigens useful for the detection of viral antigens in a test sample. The kit contains antigens obtained from self-assembled, non-infectious, non-replicating, immunogenic retrovirus-like particles comprising modified HIV-1 genomes devoid of long terminal repeats and containing nucleotide sequences encoding chimeric envelope glycoproteins. One preferred embodiment discloses the engineering of a series of expression vectors in which a synthetic oligomer encoding gp120 residues 306 to 328 (amino acids YNKRKRIHIGP GRAFYTTKNIIG) from the V3 loop of the MN viral isolate was inserted at various positions within the endogenous HIV-1.sub.LAI env gene. Expression studies revealed that insertion of the heterologous V3(MN) loop				X	

	TITLE and ABSTRACT of the patent	YES	PROBABLY YES	PROBABL Y NO	NO	I DON'T KNOW
	segment resulted in the secretion of fully assembled HIV-like particles containing chimeric LAI/MN envelope glycoproteins. Both V3 loop epitopes were recognized by loop-specific neutralizing antibodies. Immunization with HIV-like particles containing chimeric envelope proteins induced specific antibody responses against both the autologous and heterologous V3 loop epitopes, including cross-neutralizing antibodies against the HIV-1.sub.LAI and HIV-1.sub.MN isolates.					
5	**Method of casting a porous membrane of polymeric material** *Abstract* A method of casting porous membranes of polymeric material is provided wherein the membrane has at least one of the following: a more homogeneous active skin *layer for separation purposes with a lower molecular weight cut off, a more* homogeneous active skin layer with a sharper molecular weight cut off curve, a relatively thinner active skin layer with a relatively greater flux for a given separation of a substance from a solution thereof, and relatively fewer imperfections through the active skin layer. The method comprises casting a layer of a casting solution on to a support, to provide a cast layer of the casting solution thereon with an exposed skin, then progressively submerging the cast layer in a gelation liquid at a relative velocity of at least 15 cm/second, whereby the exposed skin is gently scrubbed with gelation liquid to continuously remove outwardly diffusing solvent therefrom and dissipate exothermic heat, thus causing gelation of at least a stable, more homogeneous active skin layer, with build-up of a solvent boundary layer between the exposed surface of the layer and the gelation liquid minimized.			x		
6	**Methods of treating inflammation and compositions thereof** *Abstract* Compositions and methods for treating inflammatory cell infiltration in a tissue of a mammalian subject are provided. The method involves administering a therapeutically effective amount of SERP-1, SERP-1 analog or biologically active fragment thereof admixed with a pharmaceutically acceptable carrier to a subject in				x	

	TITLE and ABSTRACT of the patent	YES	PROBABLY YES	PROBABLY NO	NO	I DON'T KNOW
	need of such treatment. Biologically active SERP-1 analogs are also provided. The compositions and methods of the present invention are useful for treating numerous inflammatory based diseases and injuries.					
7	**Vesiculated polymer granules** *Abstract* Vesiculated granules of cross-linked carboxylated polyester resin prepared by the free radical polymerization of a dispersion of an ethylenically unsaturated monomer, a carboxylated unsaturated polyester resin having an acid value of from 5-50 mg KOH/g and an alkyl acryloyl derivative. Aqueous compositions comprising the granules are used in conjunction with fibrous cellulosic or non-cellulosic materials for the preparation of paper containing the vesiculated polymeric granules dispersed therein. The paper has improved optical opacity.				x	
8	**Synthesis of long wavelength absorbing photosensitizers** *Abstract* The present invention provides for novel therapeutic macrocycle compounds useful in photodynamic therapy that are based on the chlorin ring system. The macrocycle compounds have, in many cases, wavelengths of activation at about 670 nm, characteristics of chlorins, and are stabilized against oxidation by the attachment to the chlorin ring of a structure that comprises one or more exocyclic rings that contribute at least one nitrogen atom. Protonation or covalent modification of this nitrogen atom, or other covalent modification of the one or more exocyclic rings permits optimization of pharmacalogically relevant properties including, for example, solubility. Representative chlorins include the pyridochlorins depicted as follows ##STR1##				x	
9	**Method of manufacturing a crystalline silica/platinum catalyst structure** *Abstract* A crystalline, silica platinum catalyst structure is provided by soaking a hydrophobic, high surface area, crystalline silica (SiO.sub.2) lattice essentially free of aluminum oxide in the SiO.sub.2 lattice, such as silicalite in an aqueous solution			x		

	TITLE and ABSTRACT of the patent	YES	PROBABLY YES	PROBABLY NO	NO	I DON'T KNOW
	of Pt(NH.sub.3).sub.4 Cl.sub.2 and then, drying, reducing and cooling the crystalline silica to leave, catalyst crystallites of platinum on the crystalline silica and in the pores thereof, and then providing an outer, porous, membrane coating of high molecular weight, organic polytetrafluoroethylene, polymeric material on the coated crystallites. The platinized, crystalline silica may be slurried with, polytetrafluoroethylene and a ceramic support coated with the slurry, so that a water repellent, water vapor and hydrogen or oxygen gas permeable, polytetrafluoroethylene matrix is provided on the support with the platinized, crystalline silica dispersed in the matrix. The catalyst has a high platinum content and high platinum dispersion and is particularly useful for hydrogen isotope exchange between liquid water and gaseous hydrogen or for such exothermic reactions as combining hydrogen and oxygen.					
10	**Methods for cell mobilization using in vivo treatment with hyaluronan (HA)** *Abstract* The use of forms of hyaluronic acid having a molecular weight less than about 750,000 daltons selected from the group consisting of hyaluronic acid and pharmaceutically acceptable salts thereof is provided for the same purposes known for using recombinant GM-CSF or G-CSF.					
11	**Mouse lacking the expression of interferon regulatory factor 2 (IRF-2)** *Abstract* The transcription factors, IRF-1 and IRF-2 are induced by interferons (IFNs) and a variety of other cytokines. IRF-1 functions as an activator whereas IRF-2 represses IRF-1 action by competing for binding to the same cis-elements. Recently, it has been shown that balanced expression between these two factors is critical for maintaining normal restraints on cell growth. Mutant mice deficient for IRF-2 were prepared by homologous recombination. In mutant cells, infection by Newcastle disease virus (NDV) resulted in the induction of type I IFN (IFN-.alpha. and IFN-.beta.) mRNAs, the levels of which were significantly higher than in wild type cells; whereas, such a difference was not found upon induction by poly(I):poly(C). Unlike the IRF-1 deficient mutant mice, the IRF-2 deficient mice of the invention exhibit multiple phenotypes of physical vulnerability, including lethality to lymphocytic				X	

	TITLE and ABSTRACT of the patent	YES	PROBABLY YES	PROBABLY NO	NO	I DON'T KNOW
	choriomeningitis virus (LCMV). Furthermore, in vitro colony formation assays have revealed a remarkable suppression of B cell lymphopoiesis in IRF-2 deficient mice.					
12	**Method of gas doping of vacuum evaporated epitaxial silicon films** *Abstract* A new technique for the controlled incorporation of doping impurities into homoepitaxial silicon films by gas bombardment with arsine and diborane has been investigated. Hall effect and conductivity measurements have been used to show that P and N-type silicon thin films, having mobility values close to those of bulk material may be evaporated at 700.degree.C onto silicon <111> substrates. The substrates are precleaned by evaporating 50A of silicon at 775.degree.C. Step etch measurements have shown that the gas doped films possess uniform impurity concentrations. Diode structures formed by evaporating P and N-type layers onto substrates of the opposite doping type have been investigated. These diodes show good rectification properties. Capacitance-voltage measurements confirm the abrupt nature of the evaporated junctions. Minority carrier lifetimes of 2 to 3.mu. sec. have been measured in the evaporated structures.				x	
13	**Electron beam current measuring device** *Abstract* An electron linear accelerator for use in industrial material processing, comprises an elongated, resonant, electron accelerator structure defining a linear electron flow path and having an electron injection end and an electron exit end, an electron gun at the injection end for producing and delivering one or more streams of electrons to the electron injection end of the structure during pulses of predetermined length and of predetermined repetition rate, the structure being comprised of a plurality of axially coupled resonant microwave cavities operating in the .pi./2 mode and including a graded-.beta. capture section at the injection end of				x	

	TITLE and ABSTRACT of the patent	YES	PROBABLY YES	PROBABLY NO	NO	I DON'T KNOW
	the structure for receiving and accelerating electrons in the one or more streams of electrons, a .beta.=1 section exit section at the end of the structure remote from the capture section for discharging accelerated streams of electrons from the structure and an rf coupling section intermediate the capture section and the exit section for coupling rf energy into the structure, an rf system including an rf source for converting electrical power to rf power and a transmission conduit for delivering rf power to the coupling section of the structure, a scan magnet disposed at the exit end of the structure for receiving the electron beam and scanning the beam over a predetermined product area and a controller for controlling the scanning magnet and synchronously energizing the electron gun and the rf source during the pulses.					
14	**Engraftable human neural stem cells** *Abstract* Stable clones of neural stem cells (NSCs) have been isolated from the human fetal telencephalon. In vitro, these self-renewing clones (affirmed by retroviral insertion site) can spontaneously give rise to all 3 fundamental neural cell types (neurons, oligodendrocytes, astrocytes). Following transplantation into germinal zones of the developing newborn mouse brain, they, like their rodent counterparts, can participate in aspects of normal development, including migration along well-established migratory pathways to disseminated CNS regions, differentiation into multiple developmentally- and regionally-appropriate cell types in response to microenvironmental cues, and non-disruptive, non-tumorigenic interspersion with host progenitors and their progeny. Readily genetically engineered prior to transplantation, human NSCs are capable of expressing foreign transgenes in vivo in these disseminated locations. Further supporting their potential for gene therapeutic applications, the secretory products from these NSCs can cross-correct a prototypical genetic metabolic defect in abnormal neurons and glia in vitro as effectively as do murine NSCs. Finally, human cells appear capable of replacing specific deficient neuronal populations in a mouse model of neurodegeneration and impaired development, much as murine NSCs could. Human NSCs may be propagated by a variety of means--both epigenetic (e.g., chronic mitogen exposure) and genetic (transduction of the propagating gene vmyc)--that are				X	

	TITLE and ABSTRACT of the patent	YES	PROBABLY YES	PROBABL Y NO	NO	I DON'T KNOW
	comparably safe (vmyc is constitutively downregulated by normal developmental mechanisms and environmental cues) and effective in yielding engraftable, migratory clones, suggesting that investigators may choose the propagation technique that best serves the demands of a particular research or clinical problem All clones can be cryopreserved and transplanted into multiple hosts in multiple settings.					
15	**Stabilized drill tube** *Abstract* A stabilized drill string component includes an elongated tubular body. This body comprises a tubular stabilization section and at least one further tubular section. These tubular sections are integrally joined together in axially aligned relation. The stabilization section is much shorter than the further tubular section or sections but has a somewhat greater diameter than the latter such that the stabilization section can assist in stabilizing the drill string component during drilling by way of contact with the wall of the well bore. The wearing surfaces of the stabilization section are of substantially harder material than that of the further tubular section or sections thereby to provide substantial resistance to wear.				X	
16	**Tunnel boring machine with crusher** *Abstract* The present invention relates to a tunnel boring machine particularly a micro-tunnelling machine having a cutting head rotatably mounted on the end of a housing and being driven by a first motor means and a rotatable central auger for removal of soil from the cutting operation mounted in the interior of the housing, the central auger being driven by a second motor means to allow rotation of the auger independent of the rotation of the cutting head. The tunnel boring machine may also be provided with a rock crusher between the cutting head and the auger to reduce boulders encountered during the tunneling to a size to be able to be transported by the auger.				X	

	TITLE and ABSTRACT of the patent	YES	PROBABLY YES	PROBABLY NO	NO	I DON'T KNOW
17	**Gene encoding invasion protein of campylobacter species** *Abstract* A protein associated with adherence and invasion of Campylobacter spp. including C. jejuni and C. coli is provided. Methods are disclosed for detecting Campylobacter spp. including C. jejuni and C. coli in a biological sample by determining the presence of the protein or a nucleic acid molecule encoding the protein in the sample. Compositions for treatment of infections diseases and vaccines are also described.				x	
18	**Glycosylated protein-liposome conjugates and methods for their preparation** *Abstract* The present invention provides glycosylated protein-liposome compositions which are useful for the targeted delivery of a therapeutic agent. The compositions contain an oxidized protein, typically an antibody, which is covalently attached to a lipid by means of a crosslinking agent having an acid hydrazide functionality on one terminus and a sulfhydryl functionality on the other terminus. The lipid is present in a liposome formulation. Methods for preparing the compositions are also provided. In the methods, a glycosylated protein is first oxidized then reacted with a lining group having an acid hydrazide on one end and a sulfhydryl or protected sulfhydryl group on the other end. The resultant modified protein is then reacted with a liposome formulation of a lipid having a sulfhydryl reactive functional group to covalently attach the protein to the liposome.				x	
19	**Quantum amplifier having passive core and active cladding providing signal gain by interaction of evanescent-wave components of signal and pump beams propagating along the core** *Abstract* An optical amplifier particularly for use in a high data-rate fibre-optics				x	

	TITLE and ABSTRACT of the patent	YES	PROBABLY YES	PROBABLY NO	NO	I DON'T KNOW
	communication link utilizing a body of a dielectric substance transparent to signal light of a given band of wavelengths and transparent also to pumping light of a different band of wavelengths, a smooth continuous surface along the length of the body having as amplifying component a layer of laser material adhered contiguous to the surface. Both signal light and pumping light are introduced into an end face of the body, which may be a passive dielectric fibre core such as glass, or a passive dielectric slab with parallel side walls, utilising mode order converting holograms. By the phenomenon of penetration by light of both frequencies within a thin layer of the adhered material as the light beams undergo multiple total internal reflections at the interface between the body and the laser material, atoms of active material, e.g. Nd.sup..sup.+ 3 or organic dye molecules in the laser medium are optically pumped to a higher energy level, and fluoresce and are stimulated by the evanescent-wave component of the signal light to coherently emit light in a band of wavelengths correlated with the signal band of wavelengths. The amplified signal light is delivered from the other end face, and is launched through further hologram devices in a transmission mode into a succeeding fibre or slab of the link.					
20	**Kainate-binding, human CNS receptors of the EAA2 family** *Abstract* Neurotransmission by excitatory amino acids (EAAs) such as glutamate is mediated via membrane-bound surface receptors. DNA coding for one family of these receptors, of the kainate binding type of EAA receptors, has now been isolated and the receptor protein characterized. Herein described are recombinant cell lines which produce the EAA receptor as a heterologous membrane-bound product. Also described are related aspects of the invention, which are of commercial significance. Included is use of the cell lines as a tool for discovery of compounds which modulate EAA receptor simulation.				x	

Thank you very much for you effort. Your help is greatly appreciated.

APPENDIX B (Chapter 3): SUMMARY OF THE STUDIES PROVIDING EMPIRICAL EVIDENCE

Reference	*Country & period*	*Industry(number of sectors)*	*Geographic unit (class)*	*Measure*	*Level of analysis*	*Indicator used (dependent variable)*	*EXTERNALITIES*			
							Marshall		*Jacobs*	
							Measure	*Effect*	*Measure*	*Effect*
(1) Almeida (2006)	Portugal 1985-94	manufacturing 3-digit	district (concelho) (275 distr.) (class 4)	economic growth	region	adjusted wage growth	simple location quotient	positive	inverse Herfindahl index	no
						non-adjusted wage growth		positive		no
						employment growth		negative		positive
(2) Baptista & Swann (1998)	UK 1975-82	manufacturing 2-digit (10 sect.)	CSO region (11 regions) (class 1)	innovation	firm (248 firms)	number of innovations	own industry employment	positive	other industry employment	no
									Herfindahl index	no
(3) Baptista & Swann (1999)	US 1988 & UK 1991	computer industry & services (8 sect.)	US state (39 states) (class 1) UK CSO region (14 regions) (class 1)	economic growth	firm (674 US firms) (1339 UK firms)	employment growth	own industry employment	positive (industry)	Herfindahl index	no
								no (services)		
							other industry employment (diseconomies)	positive		
					region	number of new firms	matrix of entrants and attractors	no	matrix of entrants and attractors	positive
(4) Baten *et al.* (2005)	Germany 1878 - 1913	manufacturing	region (only within Baden state) (class 2)	innovation	firm (2407 firms)	number of patents granted and renewed for min 10 years (German Patent Office)	Herfindahl index	positive	urbanization index	positive
							own industry employment in innovative firms	no	other industry employment in innovative firms	no
							own industry employment in non-	no	other industry employment in non-	no

Reference	Country & period	Industry(number of sectors)	Geographic unit (class)	Measure	Level of analysis	Indicator used (dependent variable)	EXTERNALITIES			
							Marshall		Jacobs	
							Measure	Effect	Measure	Effect
							innovative firms		innovative firms	
(5) Batisse (2002)	China 1988-97	manufacturing (30 sect.)	province (29 provinces) (class 1)	produc-tivity	region	value added growth	share of the industry value added in a region relative to the national share	negative	normalized inverse Herfindahl index of industry value added	positive
(6) Beardsell & Henderson (1999)	US 1977-92	computer industry	SMA (317 areas) (class 5)	productivity	firm	plant output (plant-level production function)	own industry employment	positive (non-cor)	Herfindahl index	no (non-cor)
								no (corpor.)		no (corpor.)
		high tech (3 sect.)						positive (hi-tech)		no (hi-tech)
(7) Beaudry (2001)	UK 1988-94	aerospace industry (5-digit) (8 sect.)	cluster - region (NUTS 3) (class 2)	economic growth	region	number of new firms	own industry employment	positive	-	-
					firm (548 firms)	employment growth	own industry employment	positive	other industry employment	negative
									employment diversity Herfindahl index	no
				innovation	firm (421 firms)	number of patents (EPO)	own industry employment in innovative firms	positive	other industry employment in innovative firms	positive
									other industry employment in non-innov. firms	negative
							own industry employment in non-	negative	employment diversity Herfindahl	positive

Reference	***Country & period***	***Industry(number of sectors)***	***Geographic unit (class)***	***Measure***	***Level of analysis***	***Indicator used (dependent variable)***	***EXTERNALITIES***			
							Marshall		***Jacobs***	
							Measure	***Effect***	***Measure***	***Effect***
							innovative firms		index	
									patent diversity Herfindahl index	positive
(8) Beaudry & Breschi (2000)	UK & Italy 1988-98	manufacturing (2-digit) (15 UK sectors) (17 Ital. sectors)	region (NUTS 3) (65 UK counties) (class 2) (95 Ital. provinces) (class 2)	innovation	firm (23 872 UK firms) (37 724 Ital. firms)	number of patents (EPO)	own industry employment	positive	other industry employment	negative
									employment diversity Herfindahl index	no
									patent diversity Herfindahl index	no
							own industry employment in innovative firms	positive	other industry employment in innovative firms	no (UK) negative (Italy)
							own industry employment in non-innovative firms	negative	other industry employment in non-innovative firms	negative (UK) positive (Italy)
(9) Beaudry & Breschi (2003)	UK & Italy 1990-98	manufacturing (2-digit) (15 UK sectors) (17 Ital. sectors)	region (NUTS 3) (65 UK counties) (class 2) (95 Ital. provinces) (class 2)	innovation	firm	number of patents (EPO)	own industry employment	no	other industry employment	no
									employment diversity Herfindahl index	no

Reference	*Country & period*	*Industry(number of sectors)*	*Geographic unit (class)*	*Measure*	*Level of analysis*	*Indicator used (dependent variable)*	***EXTERNALITIES***			
							Marshall		***Jacobs***	
							Measure	***Effect***	***Measure***	***Effect***
							own industry employment in innovative firms	positive	other industry employment in innovative firms	no
							own industry employment in non-innovative firms	negative	other industry employment in non-innovative firms	negative (UK) no (Italy)
(10) Beaudry & Swann (2001)	UK	manufacturing, services & agriculture (2-digit)	region (NUTS 3) (class 2)	economic growth	firm 137816 firms	employment growth	own industry employment	positive	other industry employment	negative
									employment diversity Herfindahl index	positive
(11) Beaudry *et al.* (1999)	Belgium France Italy Netherl. Spain Portugal UK	broad range (2-digit) (17 sect.)	region (NUTS 3) (class 2)	economic growth	firm	employment growth	own industry employment	neg (Sp, Fr, Por, UK)	other industry employment	positive (Bel. Fr., It., Net.)
				innovation	firm	number of patents (EPO) (only UK and Italy)	own industry employment	positive	other industry employment	neg (UK) no (Italy)
									employment diversity Herfindahl index	no (UK) positive (Italy)
									patent diversity Herfindahl index	no
(12) Black &	US 1972,	machinery (3 sect.)	MSA (226 areas) (class 5)	productivity	firm	plant output (plant-level	based on the region's share	no (mach.)	normalized Hirsch.-	no (mach.)

Reference	*Country & period*	*Industry(number of sectors)*	*Geographic unit (class)*	*Measure*	*Level of analysis*	*Indicator used (dependent variable)*	***EXTERNALITIES***			
							Marshall		***Jacobs***	
							Measure	***Effect***	***Measure***	***Effect***
Henderson (1999)	77, 1982, 87 and 1992	high tech (3 sect.)				production function)	in national own industry employment relative to its share in total national empl.	positive (hi-tech)	Herfindahl index	no (hi-tech)
(13) Boschma *et al.* (2005)	UK 1841 - 1971	vehicle industry (16 sect.)	county (52 counties) (class 2)	economic growth	region	employment growth	simple location quotient	positive (later stage)	relative Herfindahl index	positive (early stage)
(14) Cainelli & Leoncini (2001)	Europe 1980-92	manufacturing (9 sect.)	region (NUTS 2) (89 regions) (class 1)	economic growth	region	growth rate of labour productivity	location quotient (relative to the European level)	positive	inverse Herfindahl index	no
(15) Callejon & Costa (1996)	Spain 1980 & 1991	manufacturing (mature indus.) (18 sect.)	province (NUTS 3) (50 prov.) (class 2)	economic growth	region	employment growth	simple location quotient	positive	Herfindahl index	no
							own industry employment	no	total manufacturing industry employment	positive
(16) Capello (2002)	Italy	high tech industry & services	city (only Milan) (class 5)	productivity	firm (66 firms)	firm output (firm-level production function)	weighted indices of several elements (based on the replies from questionnaires)	positive	weighted indices of several elements (based on the replies from questionnaires)	negative (small f.) positive (big fir.)
(17) Cingano & Schivardi (2004)	Italy 1991	manufacturing (10 sect.)	local labour system (784 LLSs) (class 3)	productivity	region	total factor productivity growth	simple location quotient	positive	Herfindahl index	no
				economic growth	region	employment growth		negative		negative
						wage growth		no		no
(18)	France	manufacturing	employment zone	economic	region	employment	location	no	normalized	negative

Reference	Country & period	Industry(number of sectors)	Geographic unit (class)	Measure	Level of analysis	Indicator used (dependent variable)	EXTERNALITIES			
							Marshall		Jacobs	
							Measure	Effect	Measure	Effect
Combes (2000)	1984-93	(52 sect.)	(341 zones) (class 3)	growth		growth	quotient (relative to the national share)	(manuf.)	inverse Herfindahl index	(manuf.)
		services (42 sect.)						negative (service)		positive (service)
(19) Combes *et al.* (2004)	France 1984-93	manufacturing & services (36 sect.)	employment zone (341 zones) (class 3)	economic growth	region	average plant size growth	auteregressive coefficients	no	number of active industries in a region	no
									inverse Herfindahl index	positive
						average number of plants growth		no	number of active industries in a region	no
									inverse Herfindahl index	positive
(20) Costa-Campi & Viladecans (1999)	Spain 1991	manufacturing (21 sect.)	municipality (332 munic.) (class 5)	produc-tivity	region	export over sales (capacity to export)	simple location quotient	positive	index of global specialization	no
(21) Cota (2002)	Mexico 1988-98	manufacturing (4-digit) (54 sect.)	city (class 5)	economic growth	region	employment growth	location quotient (relative to the national share)	positive (more)	squared share of own ind. employment weighted by the national index	positive (less)
									city population	positive
(22) de Lucio *et*	Spain 1978-92	manufacturing (30 sect.)	province (class 2)	economic growth	region	employment growth	location quotient	negative	Herfindahl index	positive

Reference	*Country & period*	*Industry(number of sectors)*	*Geographic unit (class)*	*Measure*	*Level of analysis*	*Indicator used (dependent variable)*	*EXTERNALITIES*			
							Marshall		*Jacobs*	
							Measure	*Effect*	*Measure*	*Effect*
al. (1996)							(relative to the national share)			
(23) de Lucio (2002)	Spain 1978-92	manufacturing (26 sect.)	province (50 prov.) (class 2)	productivity	region	average firm output	simple location quotient	negative (first) positive (later)	Herfindahl index	no
(24) Deidda *et al.* (2002)	Italy 1991-96	manufacturing & services (2-digit) (34 sect.)	local labour system (784 LLSs) (class 3)	economic growth	region	employment growth	index of relative production specialization	negative	inverse Herfindahl index	positive
(25) Dekle (2002)	Japan 1975-95	manufacturing & services (1-digit) (9 sect.)	prefecture (47 prefect.) (class 2)	productivity	region	total productivity factor growth	share of own industry output in a region relative to the national share	positive	Herfindahl index	no
(26) Duranton & Puga (2001)	France 1993-96	manufacturing & services	employment zone (341 zones) (class 3)	economic growth	region	entry & exit of firms (establishment relocations)	normalized simple location quotient	positive (later stages)	normalized inverse Herfindahl index	positive (early stages)
(27) Ejermo (2005)	Sweden 1991-99	broad range	local labour market region (class 3)	innovation	region	number of patent applications (EPO)	index based on technological closeness of patents	positive	index based on technological closeness of patents	no
(28) Feldman & Audretsch (1999)	US 1982	manufacturing (4-digit) (grouped into 6 groups)	SMA (class 5)	innovation	region	number of new inventions reported by trade journals (SBIDB)	simple location quotient	negative	share of the industries with common science base relative to the national share	positive

Reference	Country & period	Industry(number of sectors)	Geographic unit (class)	Measure	Level of analysis	Indicator used (dependent variable)	EXTERNALITIES			
							Marshall		Jacobs	
							Measure	Effect	Measure	Effect
					firm (700 firms)	number of new inventions reported by trade journals (SBIDB)	proportion of the firms innovative activity in the firm's primary industry	negative	share of innovations in common science base product categories relative to the total number of firm's innovations	positive
(29) Forni & Paba (2002)	Italy 1971-91	manufacturing (3-digit) (88 sect.)	local labour system (955 LLSs) (class 3)	economic growth	region	employment growth	specialization indices in "spillover matrix " (based on the shares of own industry employment for each industry)	positive	Herfindahl index	positive
(30) Frenken *et al.* (2005)	Netherl. 1996 - 2002	broad range (2-digit)	labour market region (COROP) (40 regions) (class 2)	economic growth	region	employment growth	Los index (technological relatedness between industrial sectors)	no	related variety (variety measured within sectors, excluding the variety effects between sectors)	positive
				produc-tivity		productivity growth		no		negative
(31) Gao (2003)	China 1985-93	manufacturing (2-digit) (32 sect.)	province (29 prov.) (class 1)	produc-tivity	region	output growth rate	location quotient (relative to the national	negative	Herfindahl index	no

Reference	Country & period	Industry(number of sectors)	Geographic unit (class)	Measure	Level of analysis	Indicator used (dependent variable)	EXTERNALITIES			
							Marshall		Jacobs	
							Measure	Effect	Measure	Effect
							share)			
(32) Glaeser *et al.* (1992)	US 1956 & 1987	manufacturing & non-manufacturing services (6 sect.)	SMA (170 cities) (class 5)	economic growth	region	employment growth	location quotient (relative to the national share)	negative	share of other industry employment (in 5 largest industries)	positive
				productivity		wage growth		no		positive
(33) Greunz (2004)	Europe 1997-98	manufacturing (16 sect.)	region (NUTS 3) (150 high density reg.) (class 5)	innovation	region	number of patents (EPO)	location quotient (relative to the national share)	positive (lo-tech) negative (hi-tech)	reciprocal Gini indices (based on the patent or employment data)	positive
									Theil index (based on inequality between technological groups)	positive
(34) Hanson (1998)	Mexico 1980-93	manufacturing (4-digit) (54 sect.)	state (32 states) (class 1)	economic growth	region	average annual relative employment growth	location quotient (relative to the national share)	negative	the ratio of squared state employment shares to squared national employment shares	no
(35) Harrison *et al.* (1996)	US 1987	metal-working (3-digit) (21 sect.)	SMA (class 5)	innovation	firm (1000 firms)	likelihood of adoption of a new technology	number of own industry plants	no	Urban to Rural Continuum Codes (representing the level of	positive
							number of own ind. plants with >	positive		

Reference	*Country & period*	*Industry(number of sectors)*	*Geographic unit (class)*	*Measure*	*Level of analysis*	*Indicator used (dependent variable)*	*EXTERNALITIES*			
							Marshall		*Jacobs*	
							Measure	*Effect*	*Measure*	*Effect*
							250 employment		urbanization)	
							number of own ind. plants with > 500 employment	no		
(36) Henderson (1986)	US (1970) & Brazil (1970)	manufacturing (2-digit) (229 sect.)	US MSA (238 areas) Brazilian urban area (126 areas) (class 5)	productivity	region	output per labour-hour	own industry employment	positive	urban area population	no
									total local employment	no
(37) Henderson *et al.* (1995)	US 1970-87	capital goods manufacturing (2-digit) (5 sect.)	SMA (224 areas) (class 5)	economic growth	region	employment growth	own industry employment	positive	Herfindahl index	no
							simple location quotient	positive		
							own industry employment over urban land area	positive		
		high tech manufacturing (2-digit) (3 sect.)	SMA (224 areas) (class 5)	economic growth	region	employment growth	own industry employment	positive	Herfindahl index	positive
							simple location quotient	no		
(38) Henderson (1997)	US 1977-90	traditional manufacturing (2-digit) (5 sect.)	MSA (742 areas) (class 5)	economic growth	region	employment growth	simple location quotient	positive (short-lived)	Herfindahl index	positive (persis-tent)
(39) Henderson	Korea 1983-93	manufacturing (23 sect.)	city (73 cities) (class 5)	produc-tivity	region	value added per	own industry employment	positive	total city population	no

Reference	*Country & period*	*Industry(number of sectors)*	*Geographic unit (class)*	*Measure*	*Level of analysis*	*Indicator used (dependent variable)*	*EXTERNALITIES*			
							Marshall		*Jacobs*	
							Measure	*Effect*	*Measure*	*Effect*
et al. 2000						production worker			diversity index (based on Hirsch.-Herf. ind.)	positive (hi-tech)
(40) Henderson (2003)	US 1972-92	high tech (3-digit) (4 sect.)	MSA (317 areas) (class 5)	profuc-tivity	firm	total productivity factor growth	number of own industry plants	positive	Ellison-Glaeser index -related measure	no
									other industry employment	no
									total manufacturing employment	no
		machinery (3-digit) (5 sect.)	MSA (317 areas) (class 5)	profuc-tivity	firm	total productivity factor growth	number of own industry plants	no	Ellison-Glaeser index -related measure	no
									other industry employment	no (non-cor) positive (corpor)
									total manufacturing employment	no (non-cor) positive (corpor)
(41) Kameyama (2003)	Japan 1972-96	manufacturing (17 sect.)	city (80 cities) (class 5)	economic growth	region	employment growth	location quotient (relative to the national share)	positive (strong)	Herfindahl index	positive (weak)
(42) Kelley & Helper	US 1991	machine-making (3-digit) (21 sect.)	BEA area (a group of counties with a city) (183	innovation	firm (314 firms)	the likelihood of adoption of a new	simple location quotient	positive	Herfindahl index	positive

Reference	Country & period	Industry(number of sectors)	Geographic unit (class)	Measure	Level of analysis	Indicator used (dependent variable)	EXTERNALITIES			
							Marshall		Jacobs	
							Measure	Effect	Measure	Effect
(1997)			areas) (class 1)			technology				
(43) King *et al.* (2002)	US 1963-97	advertising agency industry	state (51 states) (class 1)	economic growth	region	employment growth	location quotient (relative to the national share)	negative	share of employment in 5 largest industries	no
(44) Le Blanc (2000)	US 1992, 1997	IT sector manufacturing & services	state (51 states) (class 1)	economic growth	region	employment growth	location quotient (relative to the national share)	negative	convergence factor (measures the closeness of 6 main IT specializations)	positive
(45) Lee *et al.* (2005)	Korea 1981-96	traditional manufacturing	city or county (197 cities, counties or wards) (class 4)	produc-tivity	region	output growth	location quotient (relative to the national share)	neg (tradit.)	share of other industry output	positive (tradit.)
		light manufacturing (2-digit)						no (light)		no (light)
(46) Mano & Otsuka (2000)	Japan 1960-95	manufacturing (2-digit) (5 sect.)	prefecture (class 2)	economic growth	region	employment growth	own industry employment	negative	1 minus Herfindahl index	positive (weak)
(47) Massard & Riou (2002)	France 1984-96	broad range (11 sect.)	department (94 depart.) (class 2)	innovation	region	number of patents (EPO)	share of own industry RD investment in a region relative to the national share	negative	inverse normalized Herfindahl index	no
(48) Monseny (2005)	Spain 1997 - 2000	manufacturing (7 sect.)	ZIP-code region (945 reg.) (only within Catalonia) (class 4)	economic growth	region	number of new firms	own industry employment	positive (lo-tech) no (hi-tech)	total employment	positive
									inverse Herfindahl index	positive

Reference	*Country & period*	*Industry(number of sectors)*	*Geographic unit (class)*	*Measure*	*Level of analysis*	*Indicator used (dependent variable)*	*EXTERNALITIES*			
							Marshall		*Jacobs*	
							Measure	*Effect*	*Measure*	*Effect*
(49) Mukkala (2004)	Finland 1995, 1999	manufacturing (3 sect.)	region (NUTS-4) (83 mainl. regions) (class 4)	productivity	region	production function	simple location quotient	positive (strongerfor smaller frims)	urbanization index (total population in the region versus average population in regions)	no
(50) Nakamura (1985)	Japan 1979	light manufacturing (2-digit)	city (only larger cities) (class 5)	productivity	region	production function (value added over labour)	own industry output (incorporated into the parameters of production function)	positive (stronger for heavy indus.)	total city employment and urban population (incorporated into the parameters of production function)	positive (stronger for light indus.)
(51) Ouwersloot & Rietveld (2000)	Netherl. 1992	manufacturing & services (2-digit)	municipality (800 munic.) (class 4)	innovation	firm (3910 firms	R&D intensity (share of employees in research)	share of own industry in total industry	negative	1 minus Herfindahl index	positive (manuf.) no (service)
(52) Paci & Usai (1999)	Italy 1990-91	manufacturing (3-digit) (85 sect.) (grouped into 6 groups)	local labour system (292 LLSs) (only in northern regions) (class 3)	innovation	region	number of patents per capita	production specialization (location quotient relative to the national share)	positive	production diversity (inverse Gini index based on employment data)	positive (strongerfor city, hi-tech)
							science base specialization (share of the complementary	positive	science base diversity (inverse Gini index)	positive

Reference	Country & period	Industry(number of sectors)	Geographic unit (class)	Measure	Level of analysis	Indicator used (dependent variable)	EXTERNALITIES			
							Marshall		Jacobs	
							Measure	Effect	Measure	Effect
							industries)			
(53) Paci & Usai (2000)	Italy 1990-91	manufacturing (3-digit) (85 sect.)	local labour system (109 LLSs) (only innovative regions) (class 3)	innovation	region	number of patents per capita	production specialization (location quotient relative to the national share)	positive	production diversity (inverse Gini index based on employment data)	positive (strongerfor city, hi-tech)
									science base diversity (inverse Gini index)	positive
							science base specialization (share of the complementary industries)	positive	innovation diversity (inverse Gini index based on patent data)	positive
(54) Paci & Usai (2005)	Italy 1991 2001	manufacturing (21 sect.) & services (13 sect.) (2-digit)	local labour system (784 LLSs) (class 3)	economic growth	region	employment growth	location quotient (relative to the national share)	negative	inverse normalized Herfindahl index	positive
(55) Rosenthal & Strange (2003)	US 1996-97	manufacturing (6 sect.)	ZIP-code regions (class 4)	economic growth	region	number of new firms per square mile	own industry employment	positive	other industry employment	no
									inverse Herfindahl index	positive
						employment in new firms	own industry employment	positive	other industry employment	no

Reference	Country & period	Industry(number of sectors)	Geographic unit (class)	Measure	Level of analysis	Indicator used (dependent variable)	EXTERNALITIES			
							Marshall		Jacobs	
							Measure	Effect	Measure	Effect
						per square mile			inverse Herfindahl index	positive
(56) Shefel & Frenkel (1998)	Israel	high tech (electron.)	region (8 sub-areas) (only within North. reg.) (class 2)	innovation	firm (211 firms)	the rate of innovation	own industry employment	positive (hi-tech)	employment in service industries	positive (hi-tech)
		low tech (plastics, metal)						negative (lo-tech)		no (lo-tech)
(57) Suedekom & Blien (2005)	Germany 1993 - 2001	manufacturing (15 sect.)	region (NUTS 3) (438 reg.) (class 2)	economic growth	region	employment growth	location quotient (relative to the national share)	negative (manuf.)	modified Herfindahl index	positive (manuf.)
		services (10 sect.)						positive (service)		positive (service)
(58) Susiluto & Loikkanen (2001)	Finland 1988-99	broad range	region (NUTS 4) (83 reg) (class 4)	produc-tivity	region	efficiency scores (based on valued added and taxable income)	Herfindahl index	positive	Herfindahl index	no
									population size	positive
(59) Swann & Prevezer (1996)	US 1988, 1991	computing & biotech.	state (class 1)	economic growth	firm	employment growth	own industry employment	positive	other industry employment	negative
					region	number of new firms		no		positive (stronger biotech.)
(60) Usai & Paci (2003)	Italy 1991-96	manufacturing (3-digit) (97 sect.)	local labour system (784 LLSs) (class 3)	economic growth	region	employment growth	location quotient (relative to the national share)	negative	inverse normalized Herfindahl index	positive
(61) van der Panne (2004)	Netherl.	broad range (2-digit) (58 sect.)	postal code region (98 reg.) (class 4)	innovation	firm (398 firms)	number of new inventions reported by trade journals	location quotient (relative to the national share)	positive (stronger for small, R&D intensity)	inverse Gini index	no

Reference	*Country & period*	*Industry(number of sectors)*	*Geographic unit (class)*	*Measure*	*Level of analysis*	*Indicator used (dependent variable)*	*EXTERNALITIES*			
							Marshall		*Jacobs*	
							Measure	*Effect*	*Measure*	*Effect*
(62) van der Panne & van Beers (2006)	Netherl. 2000-02	broad range (2-digit) (58 sect.)	postal code region (98 reg.) (class 4)	innovation	region	number of innovators	location quotient (relative to the national share)	positive	inverse Gini index	no
					firm (398 firms)	degree of innovativeness (innovative output)		positive		no
						commercial perform. after 2 years		no		positive
(63) van Oort (2002)	Netherl.	manufacturing	municipality or Zip-code (580 munic.) (class 4)	innovation	region	R&D intensity (research labour cost per employee)	location quotient (relative to the national share)	negative	Gini index (relative to the national index)	positive
(64) van Oort & Stam (2006)	Netherl. 1996 - 2000	ICT ind. (manufacturing & services) (22 sect.)	municipality or Zip-code (580 munic.) (class 4)	economic growth	region	employment growth	location quotient (relative to the national share)	positive	Gini index (relative to the national index)	positive
						number of new firms (as a proportion of incumbent firms)		positive		positive
(65) van Soest *et al.* (2002)	Netherl. 1988-97	manufacturing & services (2-digit)	larger city (57 cities) (class 5)	economic growth	region	employment growth in larger cities	location quotient (relative to the national share)	negative	share of other industry employment (in 5 largest industries)	positive
							simple location quotient	no	share of other industry employment (in 11 largest	no

Reference	***Country & period***	***Industry(number of sectors)***	***Geographic unit (class)***	***Measure***	***Level of analysis***	***Indicator used (dependent variable)***	***EXTERNALITIES*** ***Marshall***		***Jacobs***	
							Measure	***Effect***	***Measure***	***Effect***
									industries)	
			Zip-code area (only within South Holl. province) (416 areas) (class 4)	economic growth	region	employment growth in small Zip-code areas	location quotient (relative to the provincial share)	neg	share of other industry employment (in 5 largest industries)	positive
(66) Viladecans -Marsal (2000)	Spain 1994	manufacturing (18 sect.)	municipality(larger cities + vicinity) (class 5)	economic growth	region	employment growth	employment per capita in technologically related industries	no	size of population (urbanization economies)	positive
							employment per capita in provider sectors	no	squared population (urbanization diseconomies)	positive
							own industry employment	positive		
(67) Viladecans -Marsal (2004)	Spain 1994	manufacturing (6 sect.)	city with suburb (class 5)	economic growth	region	size of employment	own industry employment	positive (tradit. sect.)	population size	positive
									total employment	positive
									diversification index	positive

APPENDIX C (Chapter 3): Industries with positive results for Marshall externalities

	Industrial sectors	*Type**
(1)	food, beverages & tobacco, basic metals, non-metallic mineral products, chemicals, textiles & clothing, wood & pulp, equipment	L M
(3)	computer industry manufacturing & services	H S
(5)	rubber, tobacco	L
(6)	computer industry	H
(7)	aerospace industry	H
(8)	textile, timber, paper, chemicals, plastics, metal goods, vehicle, transport, metal manufacturing, telecommunication, instruments, mechanical, electrical	L M H
(10)	textile, wood, chemicals, non-metallic and metal p., transport equip., furniture, construction, radio, TV, computers, transport & comm., renting, post & telecom., fin. serv., other serv.	L M H S
(11)	non-metallic minerals, chemical, instruments, timber, plastics printing	L M
(12)	high tech	H
(13)	vehicle industry	M
(14)	low tech, medium tech & high tech industries	L M H
(15)	technologically mature industries	L M
(16)	high tech industry & services	H S
(18)	water & heating, metalworking, equipment, food, clothing, pulp & paper, construction	L M
(20)	glass, pottery & ceramics, elect. & electronic mat., beverages & tobacco, chem. prod., other transport, wood & furniture, other man., machinery, textile prod., leather, leather articles & footwear	L M H
(24)	tourism (hotel & restaurant)	S
(25)	finance, services, wholesale & retail	S
(33)	low & medium tech	L M
(35)	metalworking	M
(36)	iron & steel, machinery, transport equipment, metals, chemicals, non-metallic minerals, wood, furniture, apparel, leather, food, pulp & paper	L M
(37)	machinery, electrical machinery, primary metals, transp. and instruments; high tech	L M H
(38)	metals, transportation equipment, machinery, electronics, instruments	L M H
(39)	traditional, heavy, transportation, machinery, high tech	L M H
(40)	high tech industries	H
(41)	wood & wood products, pulp & paper, plastics & plastic products, non-ferrous metals. machinery, transportation equipment	L M
(42)	machine-making	M
(48)	textiles, wood & furniture, chemical products, metal products, motor vehicles	L M
(49)	food, beverages & tobacco, wood & wood products, pulp & paper, printing & publishing, metal & metal products, machinery, electrical products, transportation equipment	L M
(50)	iron & steel, nonferrous metal, non-metallic industry, elect. ind., apparel & related prod., print. & publish., rubber & plastics, stone, clay & glass, metal prod., transport. equip.	L M
(54)	hotel & restaurant	S
(55)	software, food products, fabricated metal, machinery	L M H
(56)	electronics	H
(57)	services	S
(59)	computing & biotechnology	H
(61)	R&D intensive industries	H
(64)	information and communication technology manufacturing and services	H S
(66)	metal products, textile products, wood & furniture, food, chemical products, other means of transport	LM
(67)	textiles, leather & footwear	L

** H represents high technology, M, medium technology, L, low technology, and S, services*

APPENDIX D (Chapter 3): INDUSTRIES WITH POSITIVE RESULTS FOR JACOBS EXTERNALITIES

Ref.	***Industrial sectors***	***Type****
(1)	chemical products, textiles & clothing	L M
(3)	computer industry services	H S
(5)	beverage & food, leather & furs, logging & transport of timber and bamboo, machine building, metal and non-metal mineral products, paper-making, transportation equipment	L M
(7)	aerospace industry	H
(8)	textile, paper, chemicals, metal goods, vehicle, transport, metal manufacturing	L M
(10)	mainly services	S
(11)	chemicals, motor vehicles and instr. eng., electrical, metal manuf., transport, timber	M
(13)	vehicle industry	M
(15)	high tech industries	H
(16)	high tech industry & services	H S
(18)	services, construction industry, high tech industries	L H S
(19)	manufacturing, trade & services	S
(20)	metal prod., other minerals and derivatives, paper articles & print., glass, pottery & ceramics, elect. & electronic mat., beverages & tobacco, chem. prod., other transport, wood & furniture, other man., mach., textile products	L M H S
(24)	textiles, wearing apparel, wood products, printing & publishing, chemicals, fabricated metal products, machinery, electrical machinery, wholesale trade, medical & precision instruments, furniture, hotel & restaurant, fin. & insurance, real state activities, computers	L M H S
(28)	instruments, telecom., pharmaceuticals, el. equipment, transportation, conglomerates	M H
(33)	high tech	H
(35)	metalworking	M
(36)	non-metallic minerals, printing & publishing, furniture	L
(37)	electronic components, computers, medical equipment	H
(38)	metals, transportation equipment, machinery, electronics, instruments	M H
(39)	high tech	H
(40)	corporate machinery	M
(41)	textiles, steel, tobacco, food & beverages	L
(42)	machine-making	M
(44)	information technology manufacturing & services	H S
(45)	traditional light industries (mature and standardized), heavy industries	L M
(46)	transportation equipment, metal products	M
(48)	high tech	H
(49)	food, beverages & tobacco	L
(50)	textile & mill prod., furniture & fixtures, print. & publish., food, lumber & wood, leather products, non-electrical mach.	L M
(51)	traditional & modern manufacturing industries	L M H
(52)	high tech	H
(53)	much stronger for high tech	H
(54)	hotel & restaurant, wholesale trade, other professional services, metal products	M S
(55)	software, food products, apparel, printing & publishing, fabricated metal, machinery	L M H
(56)	electronics	H
(57)	services	S
(59)	biotechnology	H
(60)	low, medium and high tech industries	L M H
(64)	information and communication technology manufacturing and services	H S
(66)	glass, other minerals & derivatives, metal products, precision instruments and office machinery, paper articles & printing, chemical products, other means of transport, food	L M H
(67)	office & computing, chemical products, food	L M H

** H represents high technology, M, medium technology, L, low technology, and S, services*

APPENDIX E (Chapter 6): Biotechnology- Single factor ANOVA for the differences in population means

(average number of co-inventors per patent in each cluster)

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
TRT	842	2453	2.912941	2.536581
MTL	469	1515	3.229299	3.849438
VAN	258	786	3.046154	2.33763
EDM	158	518	3.277778	3.667702
CAL	129	334	2.592308	1.390638
SAS	101	332	3.288462	3.158701
WIN	34	72	2.114286	4.045378
KIN	64	171	2.676923	1.253365
OTT	286	845	2.955479	2.626877
QUE	61	173	2.833333	2.079487
HAL	20	45	2.25	1.881579
SHE	16	39	2.4375	1.595833
out	47	151	2.960784	3.158431

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	127.1085	12	10.59237	3.797459626	9.33492E-06	1.756024811
Within Groups	6998.433	2472	2.831081			
Total	7125.542	2484				

APPENDIX F (Chapter 6): NANOTECHNOLOGY- SINGLE FACTOR ANOVA FOR THE DIFFERENCES IN POPULATION MEANS

(average number of co-inventors per patent in each cluster)

SUMMARY

Groups	Count	Sum	Average	Variance
Toronto	169	486	2,87574	1,942801
Montreal	162	497	3,067901	2,287286
Ottawa	103	297	2,883495	2,20198
Vancouver	103	298	2,893204	2,410051
Edmonton	57	184	3,22807	2,929198
Quebec	23	87	3,782609	2,359684
Kingston	14	48	3,428571	1,956044
Calgary	34	116	3,411765	7,40107
out	52	136	2,615385	1,653092

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	39,00393	8	4,875491	1,989548	0,045233	1,951464
Within Groups	1734,99	708	2,450552			
Total	1773,994	716				