

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LES GRAPPES DU SECTEUR MANUFACTURIER CANADIEN: LOCALISATION ET
COLOCALISATION

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAITRISE EN ECONOMIQUE

PAR
KOUENE YODJOU CARINE LAURE

MAI 2021

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.10-2015). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

J'exprime ici toute ma gratitude à mes directeurs de mémoire Julien Martin et Kristian Behrens. Votre rigueur, vos remarques et suggestions m'ont aidé à m'améliorer. Merci pour votre disponibilité et votre compréhension, je ne pouvais espérer mieux comme encadrement.

Je suis reconnaissante aux personnes avec qui j'ai partagé mon quotidien de travail. Je pense ici à Manassé DRABO, Théophile BIEDA et William EWANE, merci pour les échanges et l'ambiance conviviale.

Mes remerciements vont également à ma famille qui m'a toujours encouragée et soutenue. Je remercie mon époux pour son inconditionnel soutien, mes deux fils Evann et Elian, qui sont ma force face aux défis de la vie.

Enfin et non des moindres, je rends grâce à Dieu pour la santé et les ressources à ma disposition qui m'ont permis de mener à terme ce projet.

DÉDICACES

À mon tendre époux.

À mes fils.

Table des matières

Listes des figures	vi
Listes des tableaux	viii
Résumé	ix
Introduction	1
Chapitre 1 Littérature sur l'identification et la colocalisation des grappes	5
1.1 Littérature sur l'identification des grappes	5
1.2 Littérature sur la coagglomération des industries	7
Chapitre 2 Données et Méthodologie	10
2.1 Données	10
2.1.1 Données pour la construction des grappes	10
2.1.2 Données pour la colocalisation	13
2.2 Méthodologie	14
2.2.1 Construction des groupes d'industries	14
2.2.2 Construction des grappes du secteur manufacturier . .	19
2.2.3 Colocalisation des grappes	22
Chapitre 3 Résultats	25
3.1 Cartographies des grappes	25
3.2 Statistiques descriptives des grappes	31
Chapitre 4 Colocalisation des grappes	37

4.1	Statistiques descriptives des données sur la colocalisation . . .	37
4.2	Analyse économétrique de la colocalisation	40
4.3	Robustesse	46
	Conclusion	49
	Annexe Chapitre A Informations additionnelles sur les données	51
	Annexe Chapitre B Géocodage	56
	Annexe Chapitre C Concordance SCIAN	59
	Annexe Chapitre D Autres statistiques et résultats	60
	Bibliographie	69

Table des figures

	Page
3.1 Cartographie des grappes en 2001 (à gauche) et 2019 (à droite)	26
3.2 Nombre de grappes selon l'emploi et le nombre d'établissements	29
3.3 Les dix grosses grappes selon l'emploi et le nombre d'établissements	31
D.1 Cartographie des grappes en 2001 (à gauche) et 2019 (à droite) : suite	60
D.2 Fonction M pour chaque groupes d'industries en 2005, r en km	65

Liste des tableaux

	Page	
2.1	Groupes industries du secteur manufacturier Canadien	18
2.2	Contribution des mesures de similarités à la formation des groupes d'industries.	19
2.3	Valeur de r en kilomètres par années et groupe d'industries	21
3.1	Effectif de grappes par année et par province	27
3.2	Loi de Zipf sur la taille des grappes en 2019	30
3.3	Les dix grosses grappes selon l'emploi et le nombre d'établissements en 2019	32
3.4	Nombre d'établissements dans les grappes par groupes d'industries et par années	33
3.5	Emploi dans les grappes par groupes d'industries et par années	34
3.6	Emploi moyen dans les grappes par années et par groupe industrie	35
4.1	Statistiques descriptives sur la distance et les forces d'agglomérations pour les grappes en 2019 et par paire d'intensité technologique	38
4.2	Corrélation entre les variables de distance et des forces d'agglomérations pour les grappes.	39
4.3	Relations entre la distance et les forces d'agglomération en 2019 : régressions OLS	41
4.4	Relations entre la distance et les d'agglomération en 2019 : régressions quantiles	43
4.5	Hétérogénéité des relations entre la distance et les forces marshalliennes en 2019 : régression quantiles	44

	Page
4.6 Hétérogénéité des relations entre la distance et les forces marshalliennes en 2019 : robustesse	47
A.1 Codes SCIAN 4-digits en 2012	52
A.2 Présentations des sources de données	55
B.1 Résultats des trois géocodages	57
B.2 Effectifs selon les critères de choix de géolocalisation	58
C.1 Choix du codes de concordance des SCIAN 6-digits entre 1997 et 2017 . . .	59
D.1 Emploi manufacturier par année et province au Canada	61
D.2 Emploi et nombre d'entreprises par groupes d'industries et par années . . .	62
D.3 Emploi manufacturier par année et secteur au Canada	63
D.4 Emploi manufacturier par année et secteur au Canada	64

RÉSUMÉ

Ce mémoire étudie la localisation et la colocalisation des grappes industrielles du secteur manufacturier canadien. Il identifie ces grappes à l'aide d'une méthode d'identification pouvant à être mise en oeuvre pour des espaces économiques assez fins. Après avoir construit les grappes pour le secteur manufacturier canadien, il considère la distance bilatérale entre les grappes comme mesure de colocalisation pour analyser les déterminants de la colocalisation des grappes. En particulier, il utilise le partage d'intrants et de sortants, le partage de main d'oeuvre et les citations de brevets pour expliquer la colocalisation des grappes.

Les grappes du secteur manufacturier canadien appartiennent à dix groupes d'industries et sont en moyenne de trois cent cinquante par an entre 2001 et 2019. Elles sont hétérogènes selon l'emploi et le nombre d'établissements dans les grappes, avec en majorité des grappes ayant peu d'entreprises et/ou peu d'emploi. La classification selon l'emploi ou le nombre d'entreprises dans les grappes permet d'obtenir trois groupes d'industries correspondant à différents niveaux d'intensité technologique à savoir faible, moyenne et haute.

Le partage de main d'oeuvre et le partage d'intrants et de sortants ont un effet positif sur la colocalisation des grappes. L'effet de ces variables varie avec le niveau d'intensité technologique des paires de grappes considérées. Lorsqu'on se restreint aux paires de grappes de faible intensité technologique ou de moyenne intensité technologique, on obtient des coefficients d'amplitude en valeur absolue plus faible, ce qui laisse penser à une moindre importance de ces déterminants pour la colocalisation de ces types de grappes.

Mots clés : Colocalisation, grappes, mécanismes d'agglomération

INTRODUCTION

Le financement des grappes industrielles est un outil de développement prisé par de nombreux gouvernements (Behrens 2016). Les programmes visant le développement des grappes industrielles ont gagné davantage de popularité dans les pays, et beaucoup de financements y sont alloués¹. Malgré cette popularité, il n'existe pas pour le Canada de description de ces grappes industrielles. L'objet de ce mémoire est alors de cartographier, de décrire et d'analyser l'ensemble des grappes industrielles canadiennes.

Ces exercices existent dans la littérature pour des pays comme les États-Unis, mais la méthode utilisée dans ce mémoire présente l'avantage d'identifier les grappes pour des espaces économiques beaucoup plus fins. Cette méthode est une adaptation des travaux de construction des grappes du secteur du textile canadien réalisé par Behrens et al. (2020), à d'autres secteurs de l'industrie manufacturière canadienne. Ce mémoire est alors intéressant à l'échelle canadienne car bien qu'ayant des économies assez interconnectées, la politique économique canadienne ne peut pas s'appuyer sur la géographie des grappes des États-Unis.

1. Le rapport de l'observatoire européen pour les clusters et le changement industriel EOCIC (2019, pg. 20), dénombre depuis 2008 environ 30 programmes de développement des grappes au sein de l'Union Européenne. Ces programmes sont à plus des trois quarts partiellement ou totalement financés par les fonds publics. Ceci correspond ainsi à des budgets annuels pouvant aller jusqu'à 144 millions d'euros. Nous avons l'exemple des "Spitzen-cluster" en Allemagne avec 600 millions d'euros alloués entre 2008 et 2017. Hors de l'Union Européenne, l'initiative des supergrappes d'innovation au Canada avec 950 millions de dollars entre 2018 et 2023, ou encore le "innovative clusters economy program" au États-Unis sont d'autres exemples.

La notion de grappe s'inspire de Porter (1990, p.16), qui l'identifie à un "*groupe géographiquement proche d'entreprises interconnectées et d'institutions associées dans un domaine particulier, liées par des points communs et des complémentarités*". Les établissements pour être constitués en grappe, doivent alors être géographiquement et industriellement proches. Des cartographies des grappes existent pour plusieurs pays², mais peu d'analyses basées sur des techniques statistiques ont été réalisées pour les grappes. Ces cartographies représentent l'intensité de l'emploi des industries ou des groupes d'industries, avec un espace géographique différent³ d'une cartographie à l'autre. Ce manque d'analyses basées sur des techniques statistiques, ainsi que les différents espaces géographiques utilisés, justifient notre intérêt pour l'étude de la localisation et la colocalisation des grappes. Nous voulons ainsi produire une cartographie des grappes, pouvant être comparable entre les économies⁴, et analyser les déterminants de la colocalisation entre les grappes. La colocalisation étudiée dans ce mémoire renvoie à la capacité pour des paires de grappes à être géographiquement proches.

Pour construire des grappes économiques, la méthodologie utilisée s'inspire de Delgado et al. (2016) et Behrens et al. (2020). Cette méthode se déploie en deux étapes. La première étape consiste à regrouper les industries du Système de Classification des Industries de l'Amérique du Nord (SCIAN) sur la base de leurs similarités technologiques. Ces similarités technologiques sont calculées de manière à refléter les mécanismes d'agglomération identifiés

2. Nous avons l'exemple du "*US Cluster Mapping Project*" pour les États-Unis, du "*European Cluster Panorama*" pour l'Europe et du "*Canadian cluster map*" pour le Canada

3. Il correspond au découpage administratif en vigueur dans le pays, ce qui varie d'une étude à l'autre, et rend par conséquent difficile la comparabilité des résultats obtenus (Delgado et al. 2016).

4. Les disques ici utilisés comme espace géographique améliorent le problème d'agrégation spatiale rencontré lorsque l'espace considéré correspond au découpage administratif habituel (Duranton 2010, p.88). En effet, ces découpages sont les provinces, les états, les zones économiques ou les régions et qui ne correspondent pas toujours à la véritable portée géographique des grappes puisque les mécanismes d'agglomération sont davantage perceptibles localement (Behrens 2016).

par Marshall (1920)⁵. La seconde étape consiste à identifier les zones de forte concentration géographique au sein de chacun de ces secteurs. Après avoir construit les grappes, nous analysons les déterminants de la colocalisation des grappes du secteur manufacturier. Nous regardons aussi l'hétérogénéité dans la colocalisation des grappes en utilisant des facteurs inspirés de ceux utilisés par Faggio et al. (2017) qui ont travaillé sur l'hétérogénéité de l'agglomération des industries. Les variables explicatives sont les proxys des mécanismes marshalliens d'agglomération préalablement construits pour ces grappes. Cette analyse des déterminants intervient après une description de la localisation et des caractéristiques d'emploi et de nombre d'établissements des grappes.

Nous obtenons dix groupes d'industries pour les codes SCIAN à 4-digits du secteur manufacturier canadien⁶. Ces groupes d'industries sont hétérogènes selon l'emploi et le nombre d'établissements dans les grappes. Cette hétérogénéité est le reflet de l'intensité technologique des groupes d'industries dont les trois catégories sont faible, moyen ou élevé. Les groupes d'industries à intensité technologique élevée tels que le transport, sont caractérisés par des grappes avec peu d'établissements de grande taille. Les groupes d'industries de faible intensité technologique comme le textile, sont caractérisés par des grappes avec plusieurs établissements de petite taille.

Les résultats économétriques montrent que le partage de main d'oeuvre et le partage d'intrants et de sortants sont les deux mécanismes qui expliquent la colocalisation des grappes industrielles. Ce résultat est en accord avec la littérature sur la coagglomération des industries dans laquelle une augmentation de la coagglomération est associée à des niveaux élevés de partage de main d'oeuvre et de partage d'intrants et de sortants. L'effet des citations

5. Ces mécanismes sont le partage de main d'oeuvre, les externalités de connaissances et le partage d'intrants et de sortants.

6. Ce sont les groupes d'industries de l'aliment, du bois, du textile, du papier, du charbon, des minéraux, du caoutchouc, du transport, des appareils et de l'électronique.

des brevets sur la colocalisation est contraire à l'effet des citations de brevets sur la coagglomération des industries. Cela signifie alors que les grappes peuvent bénéficier des externalités des brevets de grappes sans être nécessairement proches. L'effet du partage d'intrants et de sortants et du partage de main d'oeuvre est plus faible pour les paires de grappes de faible intensité technologique.

Ce travail fournit une cartographie des grappes au Canada avec une méthode qui s'affranchit du découpage administratif du Canada. Il présente également les caractéristiques de ces grappes selon l'emploi et le nombre d'établissements. Il contribue à la littérature par une discussion sur les déterminants de la distance bilatérale entre les grappes.

La suite du document est organisée en quatre chapitres et une conclusion. Le premier chapitre donne une brève revue de la littérature connexe à ce travail. Le deuxième aborde les données ainsi que la méthodologie utilisée pour délimiter les groupes d'industries, construire les grappes et analyser la colocalisation de ces grappes. Le troisième chapitre présente une cartographie et une analyse descriptive des grappes. Le quatrième chapitre présente une analyse économétrique de la colocalisation et est suivi d'une conclusion.

CHAPITRE 1

LITTÉRATURE SUR L'IDENTIFICATION ET LA COLOCALISATION DES GRAPPES

Nous abordons ici la littérature sur l'identification des grappes et sur la coagglomération des industries.

1.1 Littérature sur l'identification des grappes

Un des objets de la littérature en économie géographique est de comprendre les forces qui poussent l'activité économique à se concentrer spatialement (Krugman (1991)). Les gains à l'agglomération s'expliquent par différents mécanismes tels que les biens intermédiaires, la main d'oeuvre qualifiée, les coûts de transports faibles ou encore les avantages naturels (Marshall 1920, Porter 1990, Krugman 1991, Ellison & Glaeser 1997). L'agglomération des économies renvoie à la notion de grappes économiques qui a été popularisée dans les années 1990 par l'économiste Michael Porter. Ces grappes ont suscité l'intérêt de la recherche en économie géographique et des politiques publiques, et cet intérêt est encore d'actualité. C'est ainsi que des auteurs se sont intéressés à documenter la relation entre les grappes et les performances des économies notamment en matière d'innovation, d'entrepreneuriat, d'emploi et de résilience¹.

1. Nous avons dans ce cas les travaux de Baptista & Swann (1998), Feldman & Audretsch (1999), Delgado et al. (2010), Martin et al. (2016).

L'identification des grappes requiert au préalable une définition de la notion, et celle de Porter (1990, p16) constitue un bon point de départ. Selon cette définition, les établissements constituant la grappe doivent être géographiquement et industriellement proches. Ces deux dimensions de proximité se retrouvent dans la plupart des définitions de grappes économiques jusque-là utilisées. Cependant, les limites géographiques et industrielles de la grappe n'étant pas clairement définies, on assiste à diverses définitions et applications du concept. Cela rend alors assez souvent la définition spécifique à l'objectif recherché, et les résultats devraient être prudemment utilisés dans un cadre de politique publique (Martin & Sunley 2003). Cette diversité de définitions rend difficile la comparaison des ensembles de grappes, et entrave la conciliation des résultats obtenus (Delgado et al. 2016).

Des travaux se sont intéressés à combler ce problème en proposant des algorithmes permettant d'identifier des grappes comparables. Delgado et al. (2016) reprennent l'algorithme de cluster développé par Everitt et al. (2011) pour regrouper les industries du code SCIAN selon des mesures de similarités. Ces mesures sont définies afin de prendre en compte les mécanismes d'agglomérations documentés dans la littérature. Elles sont calculées pour les paires de codes industries à 6 chiffres du SCIAN, en utilisant les données américaines de 2009. La dimension de proximité géographique qu'ils utilisent est celle des Aires Économiques au États-Unis.

Behrens et al. (2020) utilisent un algorithme développé par Dongen (2000) pour regrouper les industries du code SCIAN à 6 chiffres². Avec les données géocodées dont ils disposent, ils considèrent des disques de rayon 15 kilomètres (km) pour la proximité géographique. Ils s'intéressent au secteur du textile, et les grappes sont des superpositions des rayons géographiques de 15 km calculés autour des établissements dit focaux. Les établissements

2. Ils utilisent des mesures de similarités de Delgado et al. (2016).

focaux sont ceux ayant au moins 5 autres établissements du textile dans un rayon de 15 km. De plus, la probabilité d'observer plus que le nombre d'établissements du textile qui se trouvent dans son rayon de 15 km est inférieure à 0.1.

La contribution de ce mémoire à cette littérature est la généralisation des travaux de Behrens et al. (2020) à l'ensemble du secteur manufacturier du Canada. Il propose aussi un choix du rayon de voisinage pour chaque industrie qui présente l'avantage de pouvoir être appliquée à d'autres économies.

1.2 Littérature sur la coagglomération des industries

Cette littérature est celle qui inspirera notre analyse de la colocalisation des grappes. En effet, la coagglomération de deux industries renvoie à la capacité des entreprises de ces deux industries à être géographiquement proche. La colocalisation des grappes que nous étudions dans ce mémoire renvoie à la capacité des groupes d'entreprises à se localiser à proximité. Ainsi, il est possible d'avoir colocalisation des grappes sans avoir coagglomération des industries. Ces deux concepts sont alors différents.

La coagglomération des industries remonte à la fin des années 1990 avec Ellison & Glaeser (1997). Ils l'utilisent pour exprimer la tendance des industries à se localiser ensemble³ (Behrens 2016, p.1298). La littérature sur la coagglomération des industries est à ce jour largement documentée. Les mécanismes d'agglomérations identifiés par Marshall (1920) constituent la base des travaux existant à ce sujet. Pour Marshall (1920), les établissements se lo-

3. Il existe d'autres mesures de coagglomération proposées par Duranton & Overman (2005), Marcon & Puech (2010), Howard et al. (2016).

calisent proches de ceux qui utilisent les mêmes travailleurs, avec qui ils ont une relation de fournisseur-client, ou qui utilisent les mêmes technologies de production. De cette littérature, il ressort que l'agglomération de l'économie naît du partage de la main d'oeuvre, du partage d'intrants et de sortants ou encore des externalités de connaissances (Helsley & Strange 1990, Glaeser 1999).

Les travaux empiriques montrent l'effet positif du partage de la main d'oeuvre, du partage d'intrants et de sortants et des externalités de connaissances sur la coagglomération des industries. Ces résultats empiriques appuient les résultats théoriques existants. Nous pouvons citer comme exemple les travaux de Ellison et al. (2010), et Howard et al. (2016). Ellison et al. (2010) calculent sur données américaines entre 1972 et 1997 leur indice de coagglomération d'Ellison et Glaeser pour l'industrie manufacturière. Ils régressent ensuite cet indice sur les proxys des trois mécanismes d'agglomération suscités. Ils montrent que le partage d'intrants et de sortants est la force d'agglomération la plus importante entre les industries du secteur manufacturier aux États-Unis. La deuxième force d'agglomération la plus importante est le partage de la main d'oeuvre. Howard et al. (2016) utilisent des données sur des entreprises manufacturières vietnamiennes en 2007, pour calculer leur indice de colocation excessive. Ils régressent par la suite cet indice sur les proxys des forces d'agglomération. Ils obtiennent que la mise en commun de la main d'oeuvre est une force d'agglomération des industries manufacturières au Vietnam.

Ces travaux supposent une homogénéité de l'effet des mécanismes identifiés par Marshall (1920) sur l'agglomération des industries. Des travaux ont documenté l'hétérogénéité de l'effet des forces d'agglomération des industries avec des facteurs qui complètent les mécanismes marshalliens. On cite ici Chinitz (1961) qui trouve que les industries dominées par des firmes de pe-

tites tailles expérimentent des effets des forces d'agglomération plus grands. Aussi, Jacobs (1969) montrent que l'effet des externalités de connaissances est plus grand pour des espaces économiques avec beaucoup de diversité. Un autre exemple est Vernon (1960) qui soutient que les industries qui expérimentent beaucoup d'entrées bénéficient plus de l'agglomération. Ces autres facteurs sont ceux qui ont inspiré les travaux de Faggio et al. (2017) qui utilisent des données de l'industrie manufacturière au Royaume-Uni. Ils considèrent l'hétérogénéité selon l'intensité technologique, le niveau d'éducation, ou encore selon l'âge des paires d'industries. Ils montrent que les raisons de coagglomération des industries dépendent des paires d'industries considérées. En particulier la force d'agglomération de partage de la main d'oeuvre est plus grande pour les paires d'industries de faible intensité technologique. Aussi, les trois forces ont des effets plus grands lorsque l'on a les paires d'industries nouvelles.

La contribution de ce mémoire à cette littérature est qu'il adapte les travaux d'identification des grappes réalisé par Behrens et al. (2020) sur le secteur textile au Canada, à d'autres secteurs de l'industrie manufacturière. De plus il étudie les déterminants de la colocalisation d'une nouvelle unité d'analyse que sont les grappes industrielles préalablement identifiées. Nous utilisons particulièrement les trois mécanismes identifiés comme force d'agglomération des industries pour expliquer la distance bilatérale entre les grappes.

CHAPITRE 2

DONNÉES ET MÉTHODOLOGIE

Dans ce chapitre, nous allons présenter les données utilisées pour construire les grappes économiques, de même que la méthodologie de construction de ces grappes et celle pour l'étude de la colocalisation des grappes.

2.1 Données

2.1.1 Données pour la construction des grappes

Selon la définition de Porter (1990, p16), nous avons besoin des variables permettant d'avoir la proximité géographique et industrielle entre les établissements des grappes. Nous disposons ainsi d'une base sur les établissements du secteur manufacturier canadien, des données d'enquêtes et des informations sur les relations entre les paires d'industries du SCIAN. Dans ce qui suit, nous présentons chacune de ces sources de données ainsi que l'utilisation dans le calcul des mesures de proximités géographique et/ou industrielle. Les informations disponibles dans ces sources de données étant nombreuses, nous mettons l'accent uniquement sur celles d'intérêt¹.

Notre principale base de données porte sur les établissements du Canada

1. La table A.2 en annexe A présente avec plus de détail les informations disponibles dans ces sources de données.

entre 2001 à 2019, avec en moyenne 150 000 établissements par année et une périodicité de 2 ans². Elle provient de l'entreprise *Scott's Directories* qui est le principal fournisseur d'annuaires d'entreprises au Canada. Elle renseigne sur plusieurs variables telles que l'identifiant Scott de l'établissement³, son emploi, les codes SCIAN à 6-digits de son activité primaire et secondaire, son code postal et son adresse. Ces données ont été géocodées à partir des interfaces de programmations d'applications développées par Google⁴, et en utilisant l'adresse et le nom de l'établissement. Ce qui a alors permis d'obtenir des coordonnées géographiques qui permettront de représenter les établissements sur une carte du Canada.

Sur la période 2001-2019, il existe 5 versions du SCIAN correspondant respectivement aux années 1997, 2002, 2007, 2012 et 2017. D'une version à l'autre les différences observées sont la suppression, l'ajout, le regroupement ou encore la fusion de certains codes industries à 6-digits. Nous avons par conséquent réalisé une concordance des codes SCIAN afin d'avoir une stabilité sur la période d'analyse. La concordance des codes permet d'associer les établissements à un code industrie stable sur la période. Cela permet de corriger la possibilité pour deux établissements du même secteur d'activité recensés à des années différentes, d'appartenir à deux codes industries différents⁵. Le secteur manufacturier est le mieux représenté dans cette base de données Scott's. En effet, l'écart en pourcentage par rapport à l'emploi moyen entre 2001 et 2019 enregistré par Statistique Canada est plus faible pour le secteur

2. L'année 2015 n'étant pas disponible

3. Il est le même sur la période d'étude, exception faite de l'année 2001 pour laquelle nous avons utilisé le fichier de correspondance entre les identifiants Scott de 2001 et les nouveaux identifiants utilisés à partir de 2003 pour avoir des identifiants stables sur la période.

4. L'annexe B présente un détail de la procédure de géocodage.

5. Cette concordance revient à utiliser les tables de passages entre les codes industries, disponibles sur le site de Statistiques Canada <https://www.statcan.gc.ca/fra/concepts/industrie>, pour attribuer un code industrie stable à ceux ayant fait l'objet de changement au cours des années. L'annexe C présente les codes industries du secteur manufacturier du SCIAN ayant changé entre 1997 et 2017 ainsi que le code industrie choisi pour la concordance.

manufacturier (31-33), soit 15.5%⁶. Nous avons par conséquent restreint l'ensemble de données au secteur manufacturier, ce qui concerne alors 378 382 établissements géocodées entre 2001 et 2019.

Nous disposons aussi de quatre autres sources de données. Il s'agit :

- des tableaux entrées-sorties du Canada entre 1997 et 2010, qui proviennent de Statistique Canada. Elles renseignent pour chaque code SCIAN à 6-digits, la valeur des inputs en provenance de chacun des autres codes SCIAN à 6-digits ;
- de la matrice sur les citations de brevets obtenue de la base de données sur les citations du Bureau National de Recherche Economique des États-Unis. Elle donne pour chaque code SCIAN à 4-digit, la part de chacun des autres codes SCIAN à 4-digits dans ses citations de brevets ;
- des données d'enquêtes sur les occupations d'emploi, réalisées par le Bureau sur les Statistiques du Travail des États-Unis entre 2001 à 2013, avec une périodicité de 2 ans. Elles renseignent pour chaque code SCIAN à 6-digits et chaque code occupation d'emploi, le nombre d'individus occupant l'emploi relatif à ce code et dans cette industrie ;
- des données d'enquêtes courantes sur la population de 2000 à 2006, réalisées auprès des ménages par le Bureau de recensement des États-Unis. Elles renseignent pour chaque adulte dans les ménages sélectionnés, des informations telles que le code d'industrie dans lequel il exerce et le code du type d'emploi tenue.

Ces données étant à différents niveau d'agrégation pour les codes industries, nous allons considérer un niveau d'agrégation de 4-digits pour des raisons de comparaison entre les informations.

Pour caractériser la proximité industrielle entre les industries, nous allons construire cinq mesures de similarité sur la base des cinq sources de don-

6. La table D.1 en annexe A donne les informations permettant d'obtenir cet écart en pourcentage

nées à notre disposition. Les coordonnées géographiques obtenues à l'issu du géocodage de la base Scott's serviront pour la proximité géographique entre les établissements lors de la construction des grappes.

2.1.2 Données pour la colocalisation

Pour mesurer la colocalisation, nous utilisons la distance géographique entre les établissements étant aux centres des grappes. Le choix de cette mesure vient de ce que nous entendons par colocalisation la tendance pour les grappes à se localiser à proximité.

Les proxys des forces d'agglomération pour les grappes sont calculés à partir des trois forces d'agglomération des industries que sont le partage de main d'oeuvre (OES), le partage d'intrants et de sortants (IO) et les externalités de connaissances (KS). La mesure IO pour deux industries représente la proportion des biens utilisés dans ces industries et provenant de l'autre. De même, La mesure OES représente la corrélation des proportions de type de travailleurs qu'elles embauchent. Enfin, La mesure KS représente la proportion des brevets d'une industrie cités par l'autre. Plus ces proportions ou corrélations sont élevées, mieux les deux industries sont liées par ces mesures. La section 2.2.1 ci-dessous présente la méthodologie d'obtention de ces proxys pour une paire d'industrie (i,j) .

Les formules pour l'obtention des proxys pour les grappes sont les suivantes :

$$\begin{aligned}
 IO_{a,b} &= \sum_i \sum_j \rho_{ia} \rho_{jb} IO_{i,j} \\
 OES_{a,b} &= \sum_i \sum_j \rho_{ia} \rho_{jb} OES_{i,j} \\
 KS_{a,b} &= \sum_i \sum_j \rho_{ia} \rho_{jb} KS_{i,j}
 \end{aligned} \tag{2.1}$$

$IO_{a,b}$ est la force du partage d'intrants et de sortants, $OES_{a,b}$ est la force du partage de main d'oeuvre et $KS_{a,b}$ est la force des externalités de connaissances entre deux grappes a et b . De même, $IO_{i,j}$ est la force du partage d'intrants et de sortants, $OES_{i,j}$ est la force du partage de main d'oeuvre et $KS_{i,j}$ est la force des externalités de connaissances entre deux industries i et j ⁷. ρ_{ia} et ρ_{jb} représentent respectivement la part de l'industrie i dans l'emploi de a et la part de l'industrie j dans l'emploi de b .

2.2 Méthodologie

L'exercice de construction des grappes se fera en deux étapes. La première consiste à identifier les groupes d'industries étant proches selon des mesures de similarité définies. La deuxième consiste à identifier les espaces géographiques de forte concentration des établissements de ces groupes d'industries.

2.2.1 Construction des groupes d'industries

La construction des groupes d'industries nécessite des mesures de similarités entre les codes industries du SCIAN. Nous construisons ainsi cinq matrices carrées symétriques, qui représentent chacune une mesure de similarité. Chacune de ces matrices s'inspire de l'un des trois mécanismes marshalliens d'agglomération. Les lignes et colonnes de ces matrices sont identifiées par les codes industries à 4-digits du SCIAN. Les valeurs des matrices qui sont comprises entre 0 et 1, permettent de mesurer la similarité entre les paires d'industries. Ainsi, un coefficient plus grand traduit une plus grande

7. Les industries i et j sont telles que i appartienne à la grappe a et j appartienne à la grappe b .

similarité des codes industries en ligne et en colonne de cette matrice. Les cinq matrices sont définies tel que dans Behrens et al. (2020) comme suit⁸ :

- **Similarité de partage d'intrants et de sortants** : Les coefficients de cette matrice sont définis par $IO(i, j) = \text{Max}(I_{ij}, I_{ji}, O_{ij}, O_{ji})$. I_{ij} est la part de l'industrie j à 4-digits dans le total des intrants de i qui provient des autres industries à 4-digits. O_{ij} est la part de production de i qui va vers j dans le total des productions qui vont vers j . Cette matrice est construite à partir des tables entrées-sorties et permet de mesurer la force d'agglomération de partage d'intrants et de sortants ;
- **Similarité en citation de brevets** : Les coefficients de cette matrice sont définis par $KS(i, j) = \text{Max}(I_{ij}, I_{ji})$. I_{ij} est la part des brevets cités par l'industrie à 4-digits i dans toutes les citations de brevets de l'industrie à 4-digits j . Cette matrice permet de mesurer la force d'agglomération d'externalités de connaissances ;
- **Similarité d'occupation d'emploi** : Les coefficients de cette matrice sont définis par $OES(i, j) = \text{Corr}(S_i, S_j)$. S_i est le vecteur des pourcentages de chacun des 554 types occupations d'emploi dans l'emploi total de l'industrie i . Cette matrice est construite en utilisant les données d'enquêtes sur l'occupation de l'emploi, et permet de mesurer la force d'agglomération de partage de main d'oeuvre.
- **Complémentarité entre firmes** : Les coefficients de cette matrice sont définis par $WFC(i, j) = \text{Max}(S_{ij}, S_{ji})$. S_{ij} est la part des établissements de l'industrie i qui reportent l'industrie j comme activité secondaire. Cette matrice est obtenue en utilisant les données Scott's, qui reportent pour chaque établissement les codes industries des activités primaires et secondaires. Elle permet de mesurer la force d'agglomération de partage d'intrants et de sortants.

8. L'annexe A donne plus de détail sur l'obtention de certaines matrices.

- **Similarité de mobilité du travail** : Les coefficients de cette matrice sont définis par $LM(i, j) = \text{Max}(S_{ij}, S_{ji})$. S_{ij} est la fraction des travailleurs qui quittent l'industrie i pour l'industrie j . Cette matrice est obtenue en utilisant les données des enquêtes courantes sur la population. Elle permet de mesurer la force d'agglomération de partage de main d'oeuvre.

Les matrices de similarités IO, OES et WFC sont calculées pour toutes les années de disponibilité de la source de données permettant de les obtenir. Leurs coefficients sont peu variables au fil des années⁹, et nous considérons alors les matrices de l'année 2009 pour identifier les groupes d'industries.

L'identification des groupes d'industries s'inspire des travaux de Behrens et al. (2020). Ils utilisent les mesures définies par les matrices suscitées pour exécuter l'algorithme de cluster de Markov (MCL) développé par Dongen (2000). Ils exécutent cet algorithme sur les 3655 paires de codes industries à 4-digits du secteur manufacturier canadien.

Dans la mise en oeuvre de l'algorithme MCL, deux éléments sont importants à vérifier : la nature creuse de la matrice de similarité et le paramètre d'inflation de l'algorithme. Ce paramètre d'inflation noté I permet de contrôler la granularité du résultat obtenu. Une valeur élevée de I correspond à une plus grande granularité du résultat. Pour satisfaire à ces deux éléments, nous avons fixé à zéro les valeurs inférieures à la valeur médiane de toutes les valeurs de la matrice. Aussi, nous avons exécutés l'algorithme avec la valeur par défaut du paramètre I qui est de 2.

Pour chaque mesure de similarité, cet algorithme crée une partition des 85 codes industries à 4-digits du SCIAN. Les groupes d'industries sont constitués des codes industries qui se trouvent dans le même ensemble de la par-

9. On obtient pour différentes paires d'années des coefficients de corrélation des valeurs des matrices en moyenne supérieurs à 0.9.

tition pour plusieurs des cinq mesures. Il convient de noter que les résultats sur les groupes d'industries sont robustes au seuil de la valeur médiane choisie pour rendre creuse les matrices. En effet, nous obtenons les mêmes résultats en considérant le sixième décile au lieu de la médiane. Pour le paramètre d'inflation, une valeur inférieure à la valeur par défaut de 2 ne permet pas de capter des groupes industries. Par ailleurs une valeur plus élevée de 4 ne change pas les résultats obtenus.

Le tableau 2.1 présente chacun des dix groupes d'industries obtenus avec les codes industries qui les composent. On observe des différences dans la contribution des mesures à la formation des groupes d'industries. Les mesures de similarités d'intrants et de sortants, ainsi que la complémentarité des firmes sont celles ayant une grande contribution à la formation de la plupart des groupes d'industries. Les trois autres mesures, à savoir similarité d'occupation d'emploi, de mobilité du travail et en citation de brevets n'ont contribué qu'à la formation de quelques groupes d'industries.

TABLE 2.1 – Groupes industries du secteur manufacturier Canadien

Noms	Code SCIAN 4-digits	Abbréviation
Fabrication d'aliments et de boissons	3111 à 3121	Aliments
Textile	3131 à 3169 3231	Textile
Fabrication de produits en bois et de meubles	3211 à 3219, 3371 à 3379	Bois
Fabrication de papier	3221 à 3222	Papier
Plastiques et caoutchouc	3261 3262 3252	Caoutchouc
Pétrole, charbon et produits chimiques	3241 3251 3253 3254 3255 3256 3259	Charbon
Fabrication de produits minéraux non métalliques	3271 à 3279	Minéraux
Fabrication de machines industrielles, moteurs, et électroménagers	3332 3336 3339 3352 3353 3359	Appareils
Fabrication de produits informatiques et électroniques	3341 à 3346	Electronique
Equipements de transport	3361 à 3369	Transport

L'annexe A donne la définition des codes industries à 4-digits du SCIAN.

Le tableau 2.2 montre la contribution de chacune des mesures de similarité à la formation des groupes d'industries. Le symbole + désigne une faible contribution, tandis que le symbole ++ désigne une grande contribution. La distinction faible contribution vs grande contribution se réfère au nombre de codes industries à 4-digits appartenant au groupe d'industries et ayant été regroupés par la mesure de similarité. Une mesure de similarité est considérée comme ayant grandement contribué à la formation du groupe d'industries si la partition qui a été obtenue avec cette mesure a regroupé une plus grande proportion des industries appartenant à ce groupe d'industries.

TABLE 2.2 – Contribution des mesures de similarités à la formation des groupes d’industries.

Groupe d’industries	OES	WFC	LM	IO	KS
Aliments	+	++	++	++	+
Textile	++	++	++	++	+
Bois		++	+	++	
Papier		++	++		
Minéraux		++		++	
Electronique		++		++	
Caoutchouc			+	++	
Charbon				++	
Appareil		++			
Transport				++	

WFC, OES, LM, IO et KS sont respectivement les mesures de similarités selon la complémentarité des industries, la similarité d’occupation de l’emploi, la similarité selon la mobilité du travail, la similarité en terme d’intrants et de sortants et la similarité en citation de brevets.

Selon le nombre de mesures ayant contribué à la formation, nous avons trois catégories de groupes d’industries. Une première composée de groupes d’industries liés par au moins trois mesures de similarités (Aliments, Textile et Bois). La seconde est constituée des groupes d’industries liés par deux mesures de similarités (Papier, Minéraux, Electronique et Caoutchouc). La dernière catégorie comprend les groupes d’industries identifiés par une seule mesure de similarité (Transport, Charbon et Appareils).

2.2.2 Construction des grappes du secteur manufacturier

Nous identifions les grappes en s’inspirant des travaux de Behrens et al. (2020). L’idée est de déterminer les établissements ayant dans leur voisinage, une forte concentration de l’activité de son groupe d’industries. Ces établissements doivent satisfaire à un critère de taille et un critère de spécialisation. Le critère de spécialisation retient les établissements dont la probabilité d’avoir dans leur voisinage un plus grand nombre d’établissements de son

groupe d'industries par rapport aux données est jugée faible. Le critère de taille permet de raffiner l'ensemble obtenu avec le précédent critère. On retient les établissements ayant un minimum d'emploi du même groupe d'industries dans le voisinage considéré. Les grappes sont ensuite les voisinages des établissements du même groupe d'industries qui se rencontrent.

Dans les deux critères, les questions de voisinage à considérer, de seuil d'emploi minimum et de seuil de probabilité à choisir se posent. Pour déterminer le rayon du disque constituant le voisinage, nous utilisons la fonction M de Marcon & Puech (2010). Cette fonction évalue la concentration ou la dispersion relative d'une industrie dans un rayon donné. On a de la concentration relative pour des valeurs de M supérieure à 1. Cela traduit une situation où on a proportionnellement plus d'employés du groupe d'industries dans un rayon donné que dans le Canada. De même, on a de la dispersion relative pour des valeurs de M inférieure à 1. Ceci correspond aux cas où l'on a proportionnellement moins d'employés du groupe d'industries dans un rayon donné relativement à tout le Canada.

Avec des simulations de Monte-Carlo, on construit des intervalles de confiance de la fonction M pour chacun de nos groupes d'industries¹⁰. Le rayon retenu est le plus petit rayon de concentration relative significative, et correspond au premier rayon significativement supérieur à 1. Cette valeur de 1 est la valeur de référence pour l'interprétation des valeurs de M . Le tableau 2.3 donne la valeur de ce rayon par années et pour chaque groupe d'industries. Le voisinage considéré pour la construction des grappes de chaque groupe d'industries est le disque ayant pour rayon la valeur médiane des rayons sur les années pour ce groupe d'industries.

10. La Figure D.2 en annexe D donne un résultat de simulations pour l'année 2005.

TABLE 2.3 – Valeur de r en kilomètres par années et groupe d'industries

Groupes d'industries	Années									τ	μ
	2001	2003	2005	2007	2009	2011	2013	2017	2019		
Appareils	17	5	5	5	5	5	4	1	4	5	5.66
Charbon	2	25	24	1	1	1	7	25	22	7	12
Electronique	14	13	18	18	17	18	18	23	8	18	16.33
Aliments	10	14	12	11	12	10	9	10	3	10	10.11
Minéraux		7	7	6		6	6	8	8	7	6.85
Papier	12	1	3	9	9	9	8	6	2	8	6.55
Caoutchouc	23	4	4	4	2	2	2	2	16	4	6.55
Textile	4	14	14	13	4	4	12	6		9	8.87
Transport	2	4	4	4	5	60	69	4	3	4	17.22
Bois	10	35	34	32	26	26	31	32	27	31	28.11

r est le premier rayon de concentration relative significativement supérieur à 1 obtenu des simulations de monte carlo, τ est la valeur médiane et μ la valeur moyenne des r sur les années pour chaque groupe d'industries.

Pour calculer la probabilité d'avoir dans le voisinage un plus grand nombre d'établissements de son groupe d'industries par rapport aux données données, nous avons besoin de la loi du nombre d'établissements du même groupe d'industries dans le voisinage de chaque établissement. Cela permet de mettre en oeuvre le critère de spécialisation. Les établissements dont la fonction de répartition en n_i est supérieure à 0.9, sont ceux qui sont retenus par le critère de spécialisation. Le critère de taille raffine cet ensemble aux établissements qui ont en plus dans leur voisinage au moins 0.1% de l'emploi total de leur groupe d'industries¹¹.

Le nombre d'établissements du même groupe d'industries dans un voisinage s'apparente à une loi hypergéométrique $H(n_i, n_i + m_i, N+M)$ pour chaque établissement i . En effet, avoir n_i établissements du même groupe d'industries est considéré comme les n_i succès d'un tirage sans remise parmi $n_i + m_i$ éléments d'une population de taille $N+M$. Ici, N est le nombre total d'établis-

11. Le choix du seuil de 0.1% pour l'emploi autour de l'entreprise est fait pour généraliser à tout le secteur manufacturier le seuil de 5 entreprises choisi par Behrens et al. (2020) pour le textile.

sements du même groupes d'industries, M le nombre total d'établissements des autres groupes d'industries et m_i le nombre d'établissements des autres groupes d'industries dans le voisinage.

2.2.3 Colocalisation des grappes

La colocalisation des grappes étudiée ici s'inspire de la littérature sur coagglomération des industries. Les déterminants de cette coagglomération ont à ce jour été largement discutés dans la littérature empirique et théorique. La relation économétrique estimée dans ces travaux (Faggio et al. 2017, Ellison et al. 2010) se présente sous la forme

$$\delta_{i,j} = \alpha + \beta X_{i,j} + \rho D_{i,j} + \varepsilon_{i,j} \quad (2.2)$$

Ici, $\delta_{i,j}$ est l'indice de coagglomération entre les paires d'industries (i,j) , $X_{i,j}$ est un ensemble de variables décrivant les potentielles raisons de coagglomération. Ce sont par exemple les trois forces d'agglomérations identifiées par Marshall (1920) telles que le partage de main d'oeuvre, le partage d'intrants et de sortants et les externalités de connaissances. $D_{i,j}$ un ensemble d'autres variables capturant les relations entre les industries i et j , et qui sont différentes des $X_{i,j}$. $\varepsilon_{i,j}$ est un terme d'erreur.

L'intuition derrière cette équation vient de ce que les décisions de locations des entreprises dépendent des externalités (Ellison et al. (2010)). En effet Ellison et al. (2010), montrent que les entreprises s'installent à proximité l'une de l'autre pour réduire les coûts de transports dans l'acquisition des intrants ou dans la livraison des produits finis. Aussi ils montrent que les économies d'échelles dont on bénéficie lorsqu'on est situé dans un grand bassin de main d'oeuvre amène les entreprises à se situer à proximité. Enfin pour favoriser

le partage d'idées, les entreprises s'installent les unes proche des autres.

Pour appliquer l'équation 2.2 aux grappes, nous avons construits des proxys pour les forces d'agglomérations en utilisant les expressions données à l'équation 2.1. Le modèle économétrique estimé pour une année donnée est alors :

$$dist_{a,b} = \alpha + \beta_1 IO_{a,b} + \beta_2 OES_{a,b} + \beta_3 KS_{a,b} + \delta_{ab} + \varepsilon_{a,b} \quad (2.3)$$

$dist_{a,b}$ est la distance géographique entre les grappes a et b , $IO_{a,b}$ est la force de partage d'intrants et de sortants, $OES_{a,b}$ est la force de partage de main d'oeuvre et $KS_{a,b}$ est la force de citation de brevets pour les grappes a et b , δ_{ab} représente les effets fixes paires de grappes, $\varepsilon_{a,b}$ est le terme d'erreur. L'effet fixe δ_{ab} est construit de manière à prendre en compte des spécificités propres aux paires de grappes. On a autant d'indicatrices que de choix possibles de deux groupes d'industries parmi les dix groupes d'industries.

Dans la littérature, les modèles économétriques estimés considèrent d'autres variables de contrôles tels que les avantages naturels. Dans ce mémoire, pour des raisons de disponibilité des informations, nous n'allons pas inclure des contrôles dans nos spécifications. Aussi, les variables explicatives considérées sont possiblement endogène. En effet, lorsque deux entreprises sont situées à proximité, cela peut améliorer les échanges d'intrants et de bien finaux, cela peut améliorer les citations de brevets et la main d'oeuvre peut s'installer à la suite de la présence d'industries qui l'utilise.

Bien que les questions d'identifications feront l'objet de perspectives futures, nous pouvons discuter sur la nature du possible biais existant. L'omission des variables de contrôles et la possibilité d'endogénéité des variables peuvent conduire à des estimés biaisés ayant des valeurs plus élevées que la vraie valeur. Cela s'explique par le fait que ces avantages naturels sont susceptibles

d'affecter positivement les mesures de partages d'intrants et de sortants ainsi que la distance bilatérale entre les grappes. En référence à la littérature, on s'attend à un coefficient négatif pour β_1 , β_2 et β_3 .

Cette spécification à l'équation 2.2 suppose que les effets de ces forces d'agglomération sont identiques pour les paires d'industries. Nous allons regarder l'hétérogénéité de ces résultats tels qu'étudiée par Faggio et al. (2017) et dont les travaux ont été inspirés par Jacobs (1969), Vernon (1960) et Chinitz (1961). Nous allons alors estimer l'équation 2.3 par moindres carrés ordinaires et par régression sur le premier décile.

Faggio et al. (2017) utilise des régressions quantiles pour montrer que l'effet moyen du partage de main d'oeuvre, du partage d'intrants et de sortants ou des externalités de connaissances est différent selon le quantile de la distribution de l'indice de coagglomération des industries. Nous choisissons le premier décile comme quantile de la régression parce que la colocalisation étudiée ici désigne la tendance des grappes à se localiser proches. Ce qui nous intéresse alors est l'effet de ces forces sur les petites distances. Aussi, Behrens (2016) a montré que la plus grande distance de coagglomération significative entre les entreprises du secteur manufacturier canadien est de 100 km, et ce 100 km est inférieur au premier décile de la distance entre les grappes.

CHAPITRE 3

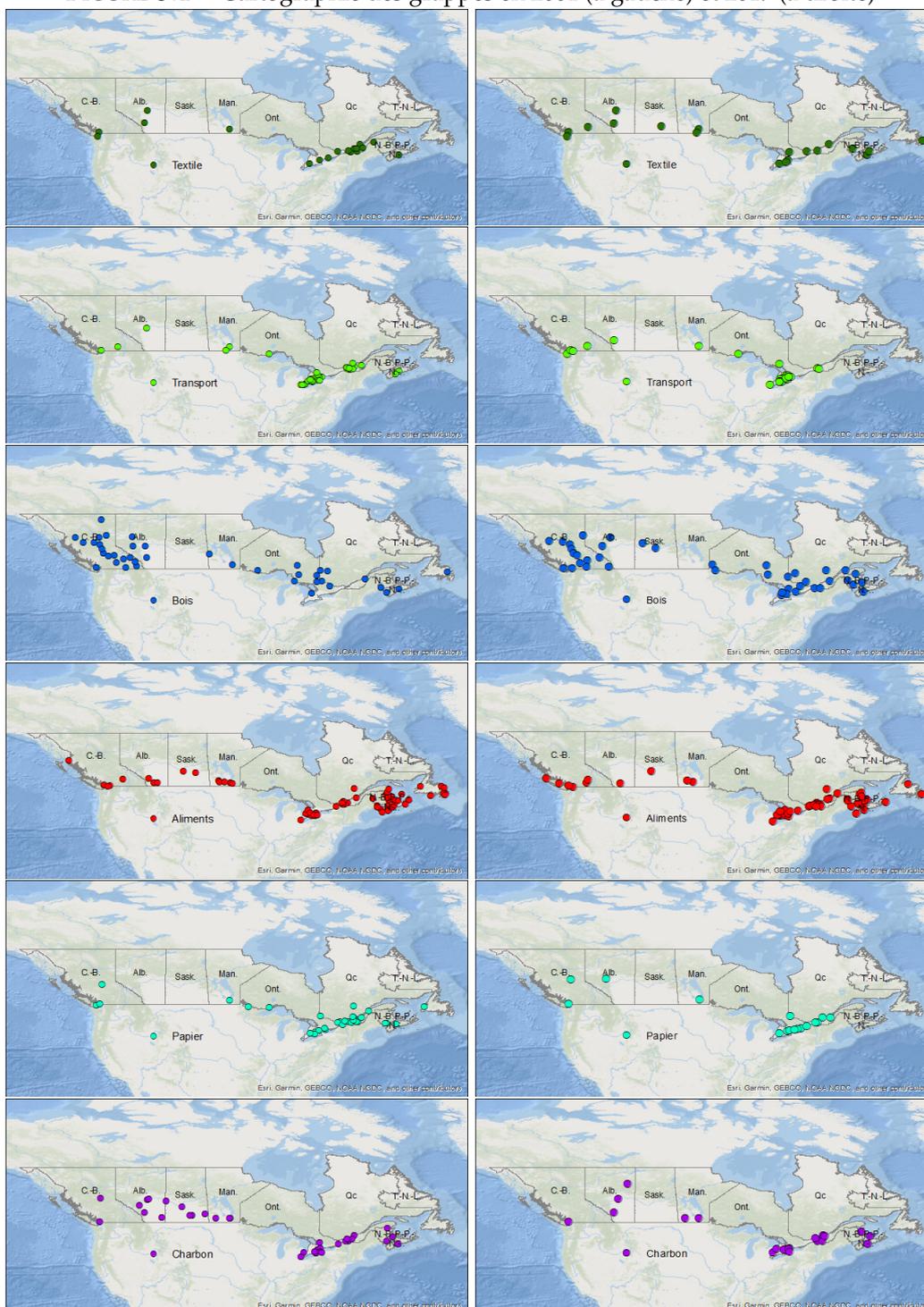
RÉSULTATS

Les résultats obtenus concernent la cartographie des grappes ainsi que quelques statistiques descriptives sur les grappes.

3.1 Cartographies des grappes

L'exercice de construction des grappes en suivant la méthodologie présentée à la section 2.2 nous a permis d'obtenir les grappes du secteur manufacturier canadien. Une carte des grappes du secteur manufacturier du Canada permet de constater qu'elles sont situées le long de la frontière canado-américaine. La localisation des grappes le long de cette frontière s'explique par la présence du marché américain qui est plus facile d'accès par cette frontière du sud du Canada. De plus, en raison des conditions climatiques au Nord du Canada, cet espace géographique constitue aussi un grand marché car la grande majorité des canadiens réside à moins de 100 km de cette frontière avec les Etats-Unis (Behrens & Bougna 2015).

FIGURE 3.1 – Cartographie des grappes en 2001 (à gauche) et 2019 (à droite)



Le panel de gauche représente l'année 2001 et celui de droite représente l'année 2019. De haut en bas, on a les grappes du groupe d'industries du Transport (en vert), du Bois (en bleu), des Aliments (en rouge), du Papier (en vert clair) et du Charbon (en violet).

La figure 3.1 ci-dessus donne une cartographie des grappes de six groupes d'industries à savoir textile, transport, bois aliments, papier et charbon en 2001 et 2019¹. On remarque que ce portrait de localisation n'a pas beaucoup évolué entre 2001 et 2019.

Entre 2001 et 2019, on observe une baisse du nombre de grappes du secteur manufacturier, cette baisse étant moins importante que celle observée pour l'emploi manufacturier. En effet le tableau 3.1 permet de constater une baisse de 15% pour le nombre de grappes qui est plus faible que la baisse de 20%² pour l'emploi manufacturier sur cette période.

Regardons la répartition du nombre de grappes obtenus par provinces entre 2001 et 2019 dans le tableau 3.1 ci dessous.

TABLE 3.1 – Effectif de grappes par année et par province

Années	Provinces											Total
	AB	BC	MAN	NB	NFL	NS	ON	PEI	QC	SASK		
2001	N	26	43	18	17	10	23	128	4	96	7	372
	%	7	12	5	5	3	6	34	1	26	2	100
2003	N	32	44	18	17	13	23	132	4	98	6	387
	%	8	11	5	4	3	6	34	1	25	2	100
2005	N	34	44	16	21	10	23	133	4	96	10	391
	%	9	11	4	5	3	6	34	1	25	3	100
2007	N	33	44	17	11	8	16	123	6	104	7	369
	%	9	12	5	3	2	4	33	2	28	2	100
2009	N	32	40	13	12	8	15	128	6	103	6	363
	%	9	11	4	3	2	4	35	2	28	2	100
2011	N	34	40	14	11	6	19	128	5	111	8	376
	%	9	11	4	3	2	5	34	1	30	2	100
2013	N	27	46	10	14	4	17	109	4	103	9	343
	%	8	13	3	4	1	5	32	1	30	3	100
2017	N	22	42	13	12	8	12	103	3	85	8	308
	%	7	14	4	4	3	4	33	1	28	3	100
2019	N	23	42	13	17	3	11	122	2	76	6	315
	%	7	13	4	5	1	3	39	1	24	2	100

N représente Le nombre des grappes.

1. La figure D.1 en annexe D présente les grappes des quatre autres groupes d'industries.
2. Voir D.1 en annexe D.

Les provinces du Québec et de l'Ontario sont le coeur de l'activité manufacturière au Canada, avec en moyenne plus de 60% de ces grappes qui y sont localisées. Cette tendance des deux provinces à regorger de la majorité du nombre de grappes se vérifie aussi au niveau de l'emploi manufacturier. Le tableau D.1 en Annexe D permet de constater qu'elles constituent plus de 70 % de l'emploi manufacturier entre 2001 et 2019.

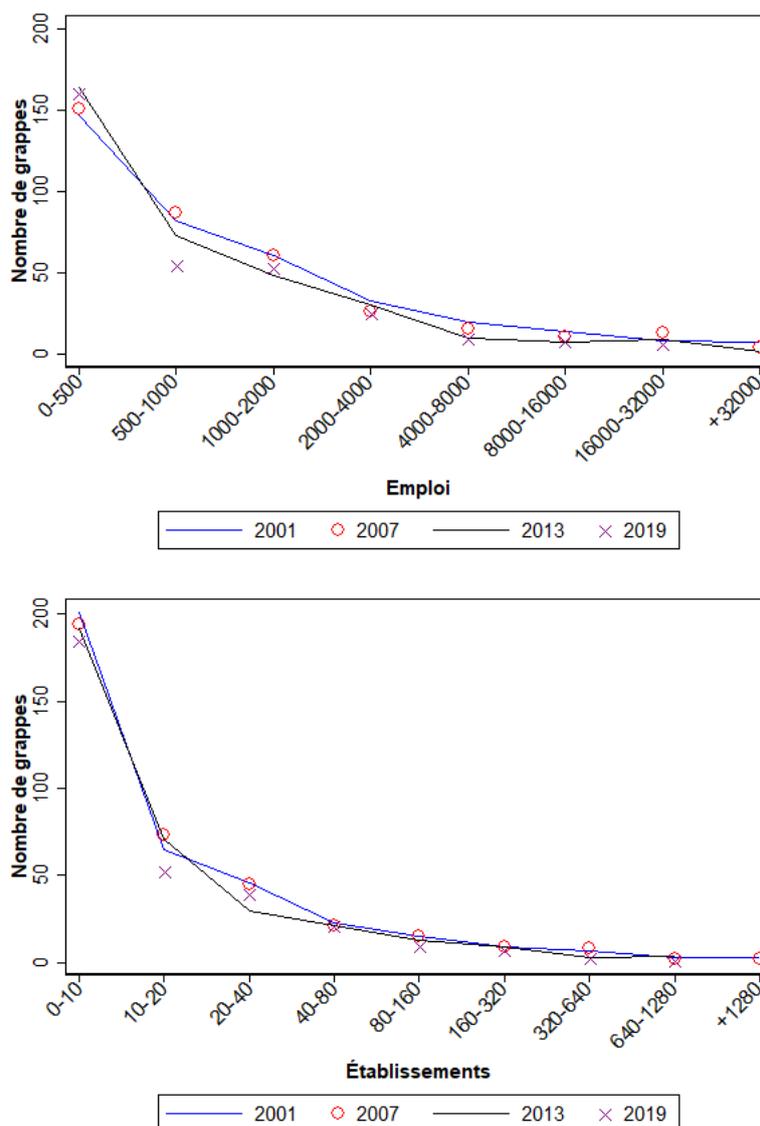
Un zoom sur la cartographie des grappes dans les dix provinces permet d'observer des spécificités. Nous avons par exemple la présence des grappes des groupes d'industries du Bois et des Aliments dans les dix provinces. Ceci serait le fait des coûts de transport élevés, ce qui incite les établissements à se localiser près des sources d'approvisionnement et de la demande finale (Roy 2010, p.15). Une autre particularité observée est la grande représentation des grappes du Bois en Colombie-Britannique. Cela s'explique par la ressource forestière disponible dans cette province et qui constitue une bonne source d'approvisionnement.

L'évolution de la taille des grappes au fil des années à la figure 3.2 permet de constater que les grappes sont en majorité constituées de peu d'établissements ou de peu d'emploi. Aussi, très peu de grappes abritent beaucoup d'établissements ou beaucoup d'emploi³. Cette tendance est restée la même de 2001 à 2019. Sur les 350 grappes en moyenne entre 2001 et 2019, plus de 200 grappes ont des emplois en dessous de 1000 ou des nombres d'établissements en dessous de 20. À l'inverse, moins de 5 grappes ont des emplois au dessus de 32000 ou des nombres d'établissements au-dessus de 1280. Cette distribution de la taille des grappes s'apparente à un fait stylisé clé de la distribution de la taille des villes, qui est la loi de Zipf. Cette loi est aujourd'hui avérée pour les villes dans nombreux pays depuis plusieurs années. Elle sti-

3. Cette observation est beaucoup liée au critère de taille de 0.1% de l'emploi fixé à la section 2.2.2. En effet, un critère plus contraignant par exemple de 10% de l'emploi aurait conduit à un nombre moins élevé de grappes et donc un ratio plus élevé du nombre de grosses grappes

pule que le nombre de villes de tailles supérieures à S est proportionnel à $\frac{1}{S}$ (Gabaix & Ioannides 2004), où S est le rang de la ville.

FIGURE 3.2 – Nombre de grappes selon l'emploi et le nombre d'établissements



Dans leurs travaux, Gabaix & Ibragimov (2011) montrent que dans une régression OLS, l'utilisation de la variable dépendante $\text{rang}_i - 0.5$ plutôt que rang_i réduit le biais de l'estimation du paramètre β .

Pour tester la validité de cette loi sur les grappes du secteur manufacturier canadien, nous avons estimé pour l'année 2019 la spécification de Gabaix & Ibragimov (2011) suivante :

$$\ln(\text{rang}_i - 0.5) = \alpha + \beta \ln \text{Taille}_i + \varepsilon_i \quad (3.1)$$

où rang_i représente le rang de la grappe i selon le critère de Taille_i ⁴. Avec des coefficients significativement différents de -1 à la table 3.2, il n'y a pas d'évidence que la taille des grappes est inversement proportionnelle à son rang. Les coefficients inférieurs à 1 en valeur absolue, signifient que la taille des grappes est plus qu'inversement proportionnelle à leur rang. En d'autres termes parmi les grappes du secteur manufacturier, on a très peu de grandes grappes et beaucoup plus de petites grappes.

TABLE 3.2 – Loi de Zipf sur la taille des grappes en 2019

Logarithme du nombre d'Établissements	-0.779 ^c (-79.45)	Logarithme de l'emploi	-0.793 ^c (-65.54)
Constante	6.477 ^c (260.37)	Constante	9.842 ^c (124.59)
N	315	Observations	315
R ²	0.95	R ²	0.93
Test d'égalité à -1 du coefficient β			
F(1,313) = 505.38		F(1,313) = 291.62	
Prob > F = 0.0000		Prob > F = 0.0000	

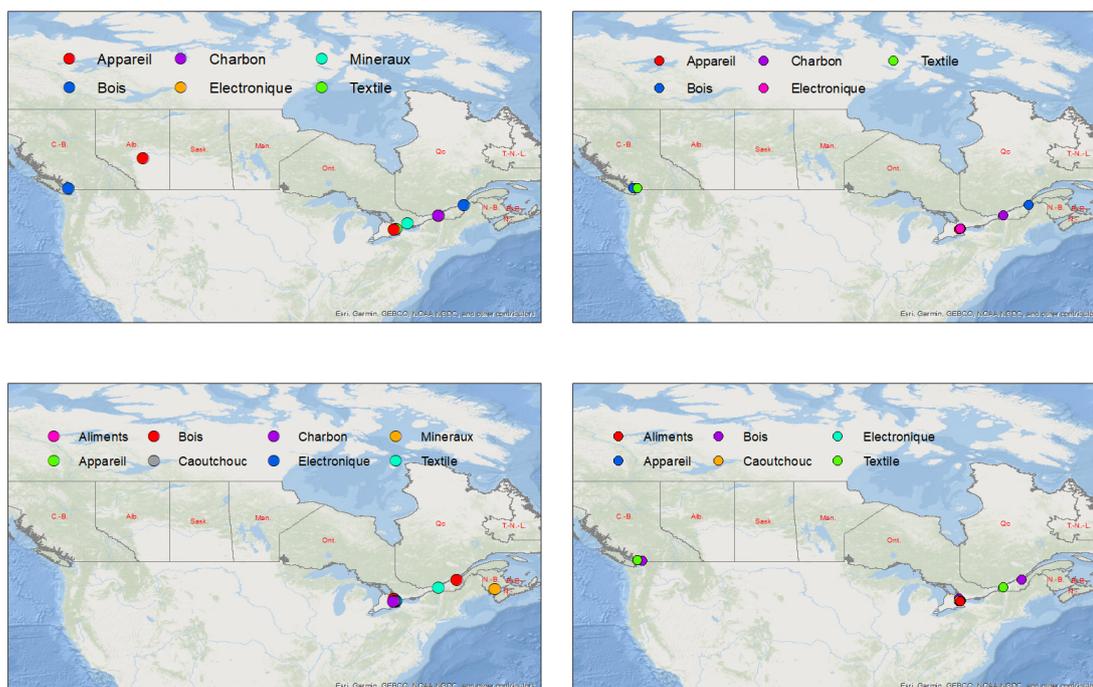
La statistique t est entre parenthèses. ^a, ^b et ^c correspondent respectivement à la significativité à 5%, 1% et 0.1%. La variable dépendante est le logarithme du rang de la grappe auquel on a soustrait 0.5.

Lorsqu'on regarde les dix plus grosses grappes à la figure 3.3, on remarque que le portrait selon l'emploi a beaucoup changé en 2019 par rapport à 2001. Les grosses grappes du Bois et des Appareils qui existaient en Colombie-Britannique et en Alberta en 2001 ne sont plus classées dans le top 10 en

4. Ici, nous avons le critère du nombre d'emploi et celui du nombre d'établissements.

2019. Ici également, le Québec et l'Ontario sont les provinces avec le plus de grosses grappes.

FIGURE 3.3 – Les dix grosses grappes selon l'emploi et le nombre d'établissements



Les grappes selon l'emploi à gauche et selon le nombre d'établissements à droite. L'année 2001 selon chaque cas se trouve au-dessus et 2019 en dessous

3.2 Statistiques descriptives des grappes

La Table 3.3 donne les caractéristiques de ces dix plus grosses grappes.

TABLE 3.3 – Les dix grosses grappes selon l’emploi et le nombre d’établissements en 2019

Grappes	Emploi	r_{emp}	N_{ets}	r_{ets}	Prov	Région Économique
Aliments	22338	1	174	10	ON	Toronto
Textile	20649	2	902	1	ON	Toronto
Textile	20580	3	284	5	QC	Montréal
Appareil	20410	4	619	2	ON	Toronto
Bois	18034	5	406	3	QC	Chaudière-Appalaches
Bois	17985	6	238	6	ON	Toronto
Electronique	15050	7	301	4	ON	Toronto
Caoutchouc	12158	8	231	7	ON	Toronto
Charbon	11003	9	x	x	ON	Toronto
Minéraux	10357	10	x	x	NB	Saint John–St. Stephen
Bois	x	x	217	8	BC	Lower Mainland–Sud-ouest
Textile	x	x	198	9	BC	Lower Mainland–Sud-ouest

x signifie que la grappe n’est pas classée parmi les 10 premières selon le critère considéré (emploi ou nombre d’établissements). N_{ets} représente le nombre d’établissements dans la grappe. Prov est la province de localisation de la grappe. r_{emp} est le rang selon le critère de l’emploi et r_{ets} le rang selon le critère du nombre d’établissements.

On constate qu’à quelques différences, les dix premières grappes selon l’emploi sont également celles selon le nombre d’établissements.

Les tableaux 3.4 et 3.5 donnent l’évolution de l’emploi et du nombre d’établissements dans les grappes par groupes d’industries entre 2001 et 2019. On remarque que les établissements de nos groupes d’industries sont restés à 50% dans les grappes, et ce malgré un déclin de 25% de l’emploi manufacturier au Canada⁵.

5. le tableau D.4 permet de calculer cette statistique

TABLE 3.4 – Nombre d'établissements dans les grappes par groupes d'industries et par années

		Aliments	Appareils	Bois	Caoutchouc	Charbon	Electronique	Mineraux	Papier	Textile	Transport	Total
2001	<i>N</i>	1.48	2.61	3.85	1.19	1.25	1.65	0.41	0.5	5.58	0.56	19.08
	%	29	58	54	48	53	80	16	56	60	34	50
2003	<i>N</i>	1.56	2.65	3.71	1.19	1.21	1.48	0.57	0.51	5.49	0.48	18.84
	%	31	58	52	46	52	81	22	58	60	29	50
2005	<i>N</i>	1.44	2.69	3.32	1.13	1.03	1.3	0.56	0.49	5.07	0.49	17.51
	%	31	56	50	44	46	80	22	58	61	30	49
2007	<i>N</i>	1.28	2.69	3.36	1.04	0.99	1.28	0.69	0.45	4.37	0.49	16.63
	%	29	58	56	43	46	80	29	57	60	31	50
2009	<i>N</i>	1.25	2.7	3.24	0.96	1	1.24	0.48	0.43	4.13	0.46	15.91
	%	29	60	57	41	47	80	20	57	60	30	50
2011	<i>N</i>	1.45	2.51	2.98	0.92	0.96	1.2	0.55	0.38	3.87	0.41	15.23
	%	34	58	57	42	48	82	25	56	60	29	50
2013	<i>N</i>	1.24	1.9	2.30	0.82	0.76	1.02	0.35	0.31	3.08	0.31	12.07
	%	35	52	55	40	44	78	19	54	59	27	48
2017	<i>N</i>	0.85	2.36	1.27	0.78	0.67	0.67	0.38	0.24	2.11	0.3	9.62
	%	29	53	45	41	47	74	22	52	58	34	45
2019	<i>N</i>	0.85	1.46	1.39	0.59	0.48	0.75	0.27	0.2	2.07	0.23	8.28
	%	33	51	50	38	39	76	21	49	59	27	46

N représente nombre d'établissements en millier

On observe par contre des différences au sein des groupes d'industries. Le groupe d'industries des Minéraux a en moyenne peu d'établissements et d'emploi dans les grappes. Par contre, les groupes d'industries du Bois, Textile, Appareil, Electronique sont les groupes d'industries caractérisés par une plus grande part d'établissements et d'emploi à l'intérieur des grappes.

TABLE 3.5 – Emploi dans les grappes par groupes d'industries et par années

		Aliments	Appareils	Bois	Caoutchouc	Charbon	Electronique	Mineraux	Papier	Textile	Transport	Total
2001	E	98.94	99.91	153.7	64.56	83.41	118.94	17.56	47.79	199.32	149.32	1033.45
	%	38	43	62	54	66	80	14	45	70	50	53
2003	E	107.85	104.08	154.37	66.93	82.47	110.32	18.82	57.98	166.3	146.82	1015.95
	%	40	65	60	52	63	88	25	44	69	61	58
2005	E	110.54	98.7	156.29	77.71	94.76	83.1	19.99	45.93	157.58	147.25	991.84
	%	41	60	60	55	65	86	28	45	70	62	58
2007	E	101.88	103.88	147.24	64.19	90.81	85.23	23.52	55.98	139.93	139.67	952.33
	%	37	63	62	50	61	88	32	52	70	60	57
2009	E	92.91	105.56	146.76	56.01	95.7	88.36	19.78	38.10	128.95	123.88	895.99
	%	35	66	63	46	63	89	19	43	70	57	55
2011	E	105.69	102.11	130.02	50.8	100.61	77.40	21.83	35.47	124.4	82.84	831.17
	%	40	62	65	47	68	89	33	45	72	41	56
2013	E	106.53	79.67	112.45	50.59	72.56	62.13	15.27	33.79	91.55	42.61	667.13
	%	44	53	62	48	55	85	24	52	69	24	50
2017	E	68.93	113.4	61.54	49.76	48.55	47.93	29.64	24.65	52.25	72.45	569.09
	%	37	59	53	52	49	83	37	51	64	51	52
2019	E	72.24	58.07	71.29	35.05	46.24	44.81	21.98	20.58	60.88	51.14	482.27
	%	43	55	60	42	50	82	39	47	61	40	51

E représente l'emploi en millier.

Ces différences entre les groupes d'industries existent lorsqu'on regarde l'emploi moyen des grappes au tableau 3.6. Les grappes des groupes d'industries du Textile, Appareil et Bois sont en majorité constituées d'établissements de peu d'employés. En revanche, les grappes des groupes d'industries du Transport et du Papier sont constituées d'établissements avec beaucoup d'employés. Ces différences reflètent le type de technologie utilisée dans ces groupes d'industries. Il convient de noter au regard du tableau D.2 que les groupes d'industries du Transport et du Papier ont en moyenne peu d'établissements dans les grappes. Par contre Textile, Appareil et Bois ont en moyenne beaucoup d'établissements au sein des grappes.

TABLE 3.6 – Emploi moyen dans les grappes par années et par groupe industrie

Année	Aliments	Appareil	Bois	Caoutchouc	Charbon	Electronique	Minéraux	Papier	Textile	Transport	Total
2001	67	38	40	54	67	72	43	96	36	265	54
2003	69	39	42	56	68	75	33	113	30	308	54
2005	77	37	47	69	92	64	36	93	31	304	57
2007	80	39	44	62	92	67	34	124	32	287	57
2009	74	39	45	59	96	71	41	88	31	268	56
2011	73	41	44	55	105	65	40	92	32	204	55
2013	86	42	49	62	96	61	44	109	30	139	55
2017	81	48	49	64	72	72	78	103	25	246	59
2019	85	40	51	59	96	60	81	105	29	223	58

Au regard de toutes ces différences présentées par les tableaux 2.2, 3.4, 3.5, 3.6, nous avons formé trois classes pour les groupes d'industries. Cette classification se rapproche de celle faite par Hatzichronoglou (1997) pour l'organisation de coopération et de développement économiques (OCDE).

En effet, Hatzichronoglou (1997) a proposé une classification des industries manufacturières selon leur intensité technologique définie comme le ratio de la recherche et développement sur la valeur ajoutée pour les Pays de l'OCDE⁶

Nous empruntons d'ailleurs ses termes pour les nommer selon l'intensité technologique des industries qui la compose. Les trois classes selon l'inten-

6. Les classifications obtenues sont : Haute technologie (Aérospatiale, Ordinateurs, machines de bureau, Electronique-Communications, Pharmaceutique). Moyenne-haute technologie (Instruments scientifiques, véhicules automobiles, machines et appareils électriques, industries chimiques, autres matériels de transport, machines non électriques). Moyenne-faible technologie (Caoutchouc et matières plastiques, construction navale, Autres industries manufacturières, métaux non ferreux, produits minéraux non métalliques, ouvrages en métaux raffinage du pétrole, métaux ferreux). Faible technologie (Papier, imprimerie et édition, textile, habillement et cuir, Alimentation, boissons et tabac, Bois et meubles.

sité technologiques sont les suivantes :

1. Faible : Aliments, Bois, Textile, Appareil
2. Moyenne : Minéraux, Papier, Caoutchouc, Electronique
3. Haute : Charbon, Transport

CHAPITRE 4

COLOCALISATION DES GRAPPES

4.1 Statistiques descriptives des données sur la co-localisation

L'exercice de construction des grappes décrit à la section 2.2 nous a permis d'obtenir des grappes économiques géolocalisées, permettant ainsi de calculer les distances bilatérales entre les grappes. Les formules à l'équation 2.1 ont permis de calculer les proxys des forces d'agglomération pour les grappes économiques¹.

Le tableau 4.1 présente quelques statistiques descriptives pour ces variables. La distance maximale entre les grappes est de 5226.7 kilomètres. Cette valeur est enregistrée pour les paires de grappes dont les industries sont de faible intensité technologique. La valeur minimale de la distance est de 87.63 mètres et est observée pour les paires de grappes dont l'intensité technologique est moyenne ou haute. Le minimum pour les citations de brevets est de 0 et est atteint pour les paires de grappes dont l'une est de faible intensité technologique. La valeur maximale des citations de brevets est de 0.784 et est enregistrée pour les paires de grappes de faible intensité technologique.

1. Les valeurs de ces variables sont par définition comprises entre 0 et 1

TABLE 4.1 – Statistiques descriptives sur la distance et les forces d’agglomérations pour les grappes en 2019 et par paire d’intensité technologique

Toutes les paires de grappes								
Variables	N	Moy	σ	Min	Max	p10	p50	p90
distance	49455	1539.9	1306.5	0.088	5226.7	138.5	1035.5	3484.1
IO	49455	0.039	0.075	0.00016	0.831	0.00273	0.01461	0.106
KS	49455	0.029	0.05	0	0.784	0.00349	0.01517	0.055
OES	49455	0.336	0.205	0.028	1	0.134	0.281	0.631
Paires d’intensité technologique Faible-Faible								
distance	12090	1706.1	1340.3	2.843	5226.7	212.5	1235.1	3629.2
IO	12090	0.052	0.097	0.00067	0.831	0.00179	0.01369	0.156
KS	12090	0.034	0.071	0	0.784	0.00277	0.01527	0.073
OES	12090	0.352	0.248	0.058	1	0.128	0.252	0.8
Paires d’intensité technologique Moyenne-Faible								
distance	16692	1553.5	1312.5	0.828	5075.9	153	1029.2	3505.9
IO	16692	0.023	0.037	0.00034	0.292	0.0028	0.01481	0.039
KS	16692	0.02	0.023	0	0.342	0.00358	0.01314	0.048
OES	16692	0.303	0.126	0.028	0.829	0.134	0.301	0.48
Paires d’intensité technologique Haute-Faible								
distance	8112	1543.5	1268	0.582	5075.8	132.5	1169.4	3408.5
IO	8112	0.013	0.014	0.00016	0.149	0.00238	0.00804	0.031
KS	8112	0.021	0.032	0	0.409	0.00279	0.01181	0.047
OES	8112	0.305	0.191	0.036	0.875	0.123	0.227	0.622
Paires d’intensité technologique Moyenne-Moyenne								
distance	5671	1377.6	1292.5	1.275	4484.8	107.2	709.7	3387.4
IO	5671	0.081	0.117	0.00122	0.822	0.00546	0.0357	0.221
KS	5671	0.064	0.078	0.00131	0.251	0.0078	0.03926	0.244
OES	5671	0.448	0.291	0.028	1	0.151	0.343	0.986
Paires d’intensité technologique Haute-Moyenne								
distance	5564	1348.2	1250	0.088	4487.9	90.2	744.7	3325.2
IO	5564	0.039	0.043	0.00068	0.206	0.00464	0.02131	0.105
KS	5564	0.021	0.014	0.00053	0.073	0.0046	0.01866	0.042
OES	5564	0.314	0.129	0.037	0.674	0.152	0.303	0.493
Paires d’intensité technologique Haute-Haute								
distance	1326	1330.4	1227.8	4.555	4487	74.7	776.7	3298.5
IO	1326	0.092	0.141	0.00037	0.691	0.00354	0.02629	0.268
KS	1326	0.035	0.042	0.00036	0.295	0.00318	0.01383	0.091
OES	1326	0.425	0.26	0.061	0.999	0.143	0.345	0.808

N représente le nombre de paires, p10, p50 et p90 respectivement le premier, le cinquième et le neuvième décile. σ l’écart type de la variable. IO, OES et KS représentent respectivement les proxys pour les relations inputs-outputs, le partage de main d’oeuvre et les externalités de connaissances.

La valeur minimale pour la force de partage de main d'oeuvre est de 0.028 et observée pour les paires de grappes dont les deux sont de moyenne intensité technologique ou encore l'une de moyenne et l'autre de faible intensité technologique. La valeur maximale de 1 est enregistrée pour les paires de même intensité technologique. Les valeurs maximales de la relation de partage d'intrants et de sortants sont relativement grandes pour les paires de même intensité technologique. Le maximum et le minimum étant observés pour les paires de faible intensité technologique. On remarque ainsi une particularité pour les paires de grappes avec au moins une de faible intensité technologique car ces sous-ensembles sont ceux pour lesquels on observe les valeurs maximales ou minimales des variables.

Regardons par la suite la corrélation existante entre les variables au tableau 4.2. La variable de distance est négativement et significativement corrélée aux variables intrants-sortants, partage de main d'oeuvre et de citations de brevets. Cette corrélation négative est plus grande avec la force de partage de main d'oeuvre.

TABLE 4.2 – Corrélation entre les variables de distance et des forces d'agglomérations pour les grappes.

	<i>Logarithme de la distance</i>	<i>Intrants-sortants</i>	<i>Citation de brevets</i>	<i>Partage de main d'oeuvre</i>
Logarithme de la distance	1			
Intrants-sortants	-0.0415 ^a	1		
Citation de brevets	-0.0156 ^a	0.2195 ^a	1	
Partage de main d'oeuvre	-0.1089 ^a	0.5989 ^a	0.3214 ^a	1

^a est significatif à 0.1%.

4.2 Analyse économétrique de la colocalisation

Pour tester de manière économétrique la relation existante entre ces variables, nous allons réaliser conjointement des régressions par moindres carrés ordinaires et par quantiles sur le premier décile. Le choix de ces spécifications est justifié en fin de section 2.2.3, et s'inspire des travaux de la littérature comme ceux de Faggio et al. (2017). En effet, nous examinons les déterminants de la colocalisation des grappes en utilisant comme variable dépendante la distance bilatérale entre les entreprises. Puisque le 1er décile de la distribution de la distance bilatérale entre nos grappes en 2019 est de 138.5 Km, supérieur à la valeur de 100 Km qui représente la plus grande distance de coagglomération significative entre les entreprises du secteur manufacturier canadien (Behrens (2016)), nous estimons des régressions quantiles sur le premier décile.

Nous allons commencer par présenter les résultats de la régression du logarithme de la distance sur le partage d'intrants et de sortants, le partage de main d'oeuvre et les citations de brevets, par moindres carrés ordinaires et ensuite par régression sur le premier décile. Après cela, nous allons regarder l'hétérogénéité des forces d'agglomération en estimant les régressions pour différents sous-ensembles selon l'intensité technologique. Nous nous intéressons à cet intensité technologique en s'inspirant de Faggio et al. (2017). En effet, Faggio et al. (2017) ont considéré différentes variables permettant de classer les industries, parmi lesquelles nous avons la classification selon le statut technologique des industries. Ils ont montrés que les effets des facteurs marshalliens de l'agglomération varient avec les classes d'appartenance des paires d'industries. Nous voulons alors tester si l'effet du partage d'intrants et de sortants, du partage de main d'oeuvre ou des citations de brevets varie selon le niveau d'intensité technologique des paires de grappes. Le tableau 4.3 présente les résultats des régressions OLS pour l'année 2019.

TABLE 4.3 – Relations entre la distance et les forces d'agglomération en 2019 : régressions OLS

	(1)	(2)	(3)	(4)
Inputs-Outputs	-1.0777 ^d (0.1229)			-0.4321 ^d (0.1282)
Citation de brevets		0.6955 ^d (0.1159)		0.7135 ^d (0.1143)
Similarité d'emploi			-1.4538 ^d (0.0607)	-1.4048 ^d (0.0624)
Constante	6.8102 ^d (0.0073)	6.7479 ^d (0.0067)	7.2572 ^d (0.0209)	7.2366 ^d (0.0212)
<i>N</i>	49455	49455	49455	49455
<i>R</i> ²	0.057	0.055	0.065	0.066

Les écarts types robustes sont entre parenthèses. significatif à ^a = 10%, ^b = 5%, ^c = 1%. La variable dépendante est le logarithme de la distance.

Les colonnes 1 à 3 donnent les résultats des régressions univariées sur le logarithme de la distance de chacune des forces d'agglomération. Il ressort des régressions que les déterminants de la colocalisation des grappes sont la relation de partage d'intrants et de sortants, la relation de partage de main d'oeuvre et la relation de citation de brevets. Les relations de partage d'intrants et de sortants et de partage de main d'oeuvre ont un effet positif et significatif sur la colocalisation des grappes. La relation en citations de brevets elle a un effet négatif et significatif sur la colocalisation des grappes. En effet, une augmentation de 0.01 dans la relation de partage d'intrants et de sortants correspond à une baisse de 1.07% de distance moyenne et donc à une augmentation de la colocalisation entre les grappes. Une augmentation similaire pour la relation de partage de main d'oeuvre correspond à une baisse de 1.45% de la distance moyenne entre les grappes, soit une augmentation de la colocalisation entre ces grappes. Cependant une augmentation de la relation en citation de brevets de 0.01 fait augmenter la distance moyenne entre les grappes de 0.69%, et donc diminue la colocalisation des grappes.

Les effets positifs du partage intrants et de sortants et du partage de main d'oeuvre sur la colocalisation sont similaires aux résultats obtenus par Ellison et al. (2010). Tandis que l'effet négatif des citations de brevets est contraire aux résultats trouvés par Ellison et al. (2010), Behrens (2016), Faggio et al. (2017). On conclut alors que les citations de brevets ne sont pas un facteur explicatif de la proximité entre les grappes. Une explication à cela serait alors que les citations de brevets sont peu coûteuses en transport, et donc il n'est pas aussi important pour deux grappes d'être proches pour pouvoir bénéficier des externalités.

À la régression multivariée de la colonne 4, les signes des coefficients sont les mêmes que ceux des régressions 1 à 3, avec des amplitudes plus faibles en valeur absolue pour la relation de partage d'intrants et de sortants, et la relation de partage de main d'oeuvre. L'amplitude de la relation de citations de brevets devient plus grande en valeur absolue. Ainsi, le partage de la main d'oeuvre et la relation d'intrants et de sortants sont les deux variables qui expliquent la proximité et donc la colocalisation des grappes. Ce résultat corrobore celui des travaux de Ellison et al. (2010), Behrens (2016) et Faggio et al. (2017).

Les résultats précédents montrent l'effet du partage d'intrants et de sortants, du partage de main d'oeuvre et des citations de brevets sur la distance moyenne entre les grappes. En s'inspirant des travaux de Faggio et al. (2017) qui ont utilisés des régressions quantiles pour montrer qu'il existe de l'hétérogénéité dans l'effet moyen des forces d'agglomérations sur la coagglomération des industries, nous allons estimer des régressions sur le premier décile de la distance bilatérale entre les grappes. Le choix du premier décile vient des travaux de Behrens (2016) qui a montré que la distance de coagglomération significative entre les entreprises du secteur manufacturier canadien est de 100 Km, laquelle valeur est inférieure au premier décile de la distribution de

la distance bilatérale entre nos grappes. Les résultats sont présentés dans le tableau 4.4.

TABLE 4.4 – Relations entre la distance et les d'agglomération en 2019 : régressions quantiles

	(1)	(2)	(3)	(4)
Inputs-outputs	-1.5922 ^c (0.2427)			-0.7625 ^c (0.2511)
Citation de brevets		1.5307 ^c (0.2241)		1.6241 ^c (0.2195)
Similarité d'emploi			-2.0727 ^c (0.1153)	-2.0030 ^c (0.1185)
Constante	5.1048 ^c (0.0173)	4.9943 ^c (0.0165)	5.7507 ^c (0.0405)	5.7134 ^c (0.0412)
<i>N</i>	49455	49455	49455	49455
<i>R</i> ²				

Les écarts types robustes sont entre parenthèses. significatif à ^a = 10%, ^b = 5%, ^c = 1%. La régression quantile pour le décile d'ordre 1. La variable dépendante est le logarithme de la distance

Avec la régression quantile sur le premier décile, les signes restent identiques à ceux de la régression OLS, mais avec des amplitudes plus grandes en valeur absolue. En effet, les citations de brevets garde toujours un signe positif et donc ne favorisent pas la colocalisation entre les grappes. Ce résultat trouve également son explication dans le fait que les citations de brevets étant moins coûteuses en transport, les grappes peuvent bénéficier des externalités sans toutes fois être à proximité. le partage d'intrants et de sortants ainsi que le partage de main d'oeuvre ont des signes négatif et sont alors les variables qui expliquent la proximité et donc colocalisation des grappes. L'interprétation des coefficients de la régression 4 nous montre qu'une augmentation de 0.01 dans le partage d'intrants et de sortants correspond à une réduction de la distance moyenne de 0.76% entre les grappes, soit une amélioration de la colocalisation des grappes. Une augmentation similaire pour la relation de partage de main d'oeuvre correspond à une réduction de 2%

de la distance moyenne entre les grappes. Ces deux premiers résultats rejoignent les travaux de Ellison et al. (2010), Behrens (2016) et Faggio et al. (2017). Par contre une augmentation similaire de la relation en citations de brevets ferait augmenter la distance moyenne entre les grappes de 1.62%. Dans le premier décile de la distribution de la distance entre les grappes, les déterminants de la colocalisation des grappes sont le partage d'intrants et de sortants et le partage de main d'oeuvre.

A l'issue de la section 3.2, nous avons présenté une classification des nos groupes d'industries selon trois niveaux d'intensité technologique. Il est intéressant pour nous de regarder comment varie l'effet du partage d'intrants et de sortants, du partage de main d'oeuvre et des citations de brevets dans la colocalisation des grappes pour chacune des paires d'intensité technologique. Ainsi, le tableau 4.5 donne les résultats des régressions sur le premier décile pour les différentes paires d'intensité technologique.

TABLE 4.5 – Hétérogénéité des relations entre la distance et les forces marshalliennes en 2019 : régression quantiles

Paires de grappes	F-F	F-M	F-H	M-M	M-H	H-H
Inputs-Outputs	-0.2184 (0.3769)	-1.5549 ^b (0.6544)	-4.4819 (3.3282)	1.6661 ^c (0.4684)	-5.8325 ^c (1.0729)	-1.6158 ^b (0.7419)
Citation de brevets	1.76914 ^c (0.2240)	-1.5179 (1.1001)	2.6907 ^c (0.6713)	7.0395 ^c (2.5674)	-13.2083 ^c (4.1425)	9.1978 ^c (2.7772)
Similarité d'emploi	-1.2915 ^c (0.2172)	-2.1006 ^c (0.2209)	-2.3278 ^c (0.2976)	-1.7552 ^c (0.3195)	-2.9403 ^c (0.4328)	-4.7300 ^c (0.6583)
Constante	5.7917 ^c (0.0744)	5.8221 ^c (0.0742)	5.7652 ^c (0.1083)	4.9658 ^c (0.1911)	6.0994 ^c (0.1475)	6.4461 ^c (0.2906)
<i>N</i>	12090	16692	8112	5671	5564	1326
<i>R</i> ²						

Les écarts types robustes sont entre parenthèses. significatif à ^a = 10%, ^b = 5%, ^c = 1%. La régression quantile pour le décile d'ordre 1. La variable dépendante est le logarithme de la distance. Dans les paires de grappes, F, M et H représente respectivement l'intensité technologique faible, moyenne et haute du groupe d'industries auquel la grappe appartient.

De ces régressions, on observe un changement dans le signe de la relation

de partage d'intrants et de sortants pour les paires de grappes dont les deux sont de moyenne intensité technologique. Ce changement de signe signifie alors qu'une augmentation de la relation de partage d'intrants et de sortants, conduit à une augmentation de la distance moyenne dans le premier décile de la distribution de la distance bilatérale entre les grappes, soit une baisse de la colocalisation. En particulier, une augmentation de 0.01 de la relation de partage d'intrants et de sortants conduit à une augmentation moyenne de 1.66% de la distance moyenne entre les paires de grappes de moyennes intensité technologique dans le premier décile. Ce précédent résultat contraire aux travaux sur la coagglomération des industries car la relation de partage d'intrants et de sortants affecte positivement la coagglomération des industries pourrait signifier que pour les paires de grappes de moyenne intensité technologique, le transport d'intrants et de sortants serait moins coûteux qu'une installation à proximité.

On observe également que le signe de la relation de citations de brevets devient négatif pour les paires de grappes dont l'une est de moyenne et l'autre de haute intensité technologique. Précisément, une augmentation de la relation en citation de brevets de 0.01 conduit à une baisse de la distance bilatérale moyenne de 13.20% dans le premier décile de la distribution de la distance entre ces paires de grappes. Cela suggère alors que l'effet contraire aux résultats existant ne se vérifie pas pour ce type de paire de grappes. Une explication plausible à ce résultat serait alors que pour ces paires de grappes, pouvoir bénéficier efficacement des brevets cités requiert alors d'être relativement proches. Ainsi le signe positif pour la citation de brevets observés dans les régressions précédentes serait alors le fait de la prédominance des grappes de faible intensité technologique dans notre échantillon.

La relation de partage de main d'oeuvre reste celle qui affecte négativement la distance moyenne sur le premier décile de la distribution pour toutes les

paires d'intensité technologique. Elle affecte alors positivement la colocalisation des grappes des différentes paires d'industries. De plus lorsqu'on compare l'amplitude de l'effet de ces variables avec les résultats du tableau 4.4, on constate que pour les paires de grappes dont les deux sont de faible intensité technologique ou de moyenne intensité technologique, l'amplitude des résultats est plus faible en valeur absolue². L'amplitude des coefficients est plus grande pour toutes les autres paires d'intensité technologique. Cela suggère alors que le partage d'intrants et de sortants, le partage de main d'oeuvre et les citations de brevets ont un effet relativement moindre sur la colocalisation des paires de grappes de faible intensité technologique, ou de moyenne intensité technologique. Cette relative faible importance pour les grappes dont les deux sont de faible ou de moyenne intensité technologique laisse penser que ces types de grappes ont un intérêt plus élevé à se situer à proximités des grappes d'intensité technologique différente des leurs.

4.3 Robustesse

Nous faisons une analyse de robustesse des résultats en considérant une autre classification des industries en utilisant le degré de technicité des entreprises qui la constituent. Cette classification est celle du conseil économique du Canada qui utilise le ratio d'intensité technologique pour faire le classement³. Ils définissent ce ratio comme la valeur de l'ensemble des facteurs de production de haute technicité dans l'ensemble des facteurs de production. Ces valeurs sont obtenues des tableaux d'entrées-sorties de Statistique Canada. La classification obtenue est la suivante :

1. Faible : Aliments, Bois, Charbon

2. 1,29 ou 1.75 contre 2,00 dans la régression du tableau 4.4

3. <https://www150.statcan.gc.ca/n1/pub/75-001-x/1990001/article/2290-fra.pdf>

2. Moyenne : Minéraux, Papier, Textile

3. Haute : Caoutchouc, Transport, Electronique, Appareil

Nous utilisons alors cette classification pour analyser la robustesse des résultats du tableau 4.5. Le tableau 4.6 ci-dessous affiche les résultats et on observe que le signe des coefficients sont robustes pour le partage d'intrants et de sortants et pour le partage de main d'oeuvre, avec une amplitude plus grande pour la plupart des coefficients. En effet, le coefficient positif de la relation de partage d'intrants et de sortants pour les paires de grappes d'intensité technologique moyenne demeure. Aussi, les coefficients pour la relation de partage de main d'oeuvre demeure négatif. Comme expliqué dans les résultats du tableau 4.5, une explication plausible du coefficient positif de la relation de partage d'intrants et de sortants pour les paires de grappes de moyenne intensité technologique, est que le transport d'intrants et de sortants serait moins coûteux qu'une installation à proximité.

TABLE 4.6 – Hétérogénéité des relations entre la distance et les forces marshalliennes en 2019 : robustesse

Paires de grappes	F-F	F-M	F-H	M-M	M-H	H-H
Inputs-outputs	0.1450 (0.3512)	0.7534 (0.7521)	-7.8222 ^c (0.9727)	2.2629 ^c (0.5121)	-12.6755 ^c (4.6288)	-1.9824 ^c (0.5889)
Citation de brevets	2.0002 ^c (0.2179)	-10.6695 ^c (2.8932)	-0.0950 (0.6020)	7.9600 ^c (2.8699)	-7.1002 (5.3156)	8.0576 ^c (2.0590)
Similarité d'emploi	-1.3822 ^c (0.2128)	-1.7903 ^c (0.2937)	-2.8113 ^c (0.2034)	-1.5076 ^c (0.5360)	-0.7246 (0.6621)	-2.6577 ^c (0.4353)
Constante	5.8789 ^c (0.0825)	5.8996 ^c (0.0814)	5.9461 ^c (0.0627)	5.3988 ^c (0.2369)	5.6043 ^c (0.2091)	5.9704 ^c (0.2820)
<i>N</i>	6555	9019	23932	1446	4262	4241
<i>R</i> ²						

Les écarts types robustes ont entre parenthèses. significatif à ^a = 10%, ^b = 5%, ^c = 1%. La régression quantile pour le décile d'ordre 1. La variable dépendante est le logarithme de la distance. Dans les paires de grappes, F, M et H représente respectivement l'intensité technologique faible, moyenne et haute du groupe d'industries auquel la grappe appartient.

Tous ces résultats économétriques nous permettent de conclure que les mécanismes d'agglomérations identifiés pour les industries sont également valable pour les grappes. L'effet de ces mécanismes dépend des paires d'intensité technologique considérées. Il est relativement plus faible pour les paires de grappes dont les deux sont de faible intensité technologique. Le partage d'intrants et de sortants ainsi que le partage de main d'oeuvre sont les facteurs qui favorisent la proximité des grappes industrielles, ce indépendamment du niveau d'intensité technologique considéré

CONCLUSION

Dans ce mémoire, nous avons cartographié les grappes du secteur manufacturier canadien. Nous avons aussi testé la validité des mécanismes d'agglomération des industries pour l'agglomération des grappes.

Nous avons premièrement regroupé les codes industries à 4-digits du Système de Classification Nord-Américain pour le secteur manufacturier, et cet exercice a permis d'obtenir 10 groupes d'industries. Le groupe du textile identifié est en accord avec celui identifié par Behrens et al. (2020), dont nous avons emprunté la méthodologie. Nous avons ensuite identifié les espaces de forte concentration de l'activité de ces groupes d'industries à partir des données géolocalisées sur les établissements. Il ressort de l'analyse descriptive des grappes qu'elles sont hétérogènes selon emploi et le nombre d'établissements qui s'y trouvent. Cette hétérogénéité nous a permis de classer les groupes d'industries selon trois niveaux d'intensité technologique à savoir faible, moyenne et élevée.

Nous avons ensuite estimé l'effet des forces d'agglomération de partage de main d'oeuvre, de citations de brevets et de partage d'intrants et de sortants sur la colocalisation des grappes mesurée ici par la distance bilatérale entre les grappes. Ces estimations ont été réalisées par moindres carrées ordinaires et par régressions quantiles sur le premier décile. Nous avons considéré l'année 2019, ce qui correspond à 49455 paires de grappes. Nous avons enfin es-

timés les régressions pour les six sous-ensembles de paires d'intensité technologique.

Les résultats permettent de conclure que le partage d'intrants et de sortants, et le partage de main d'oeuvre permettent d'expliquer la colocalisation des grappes. Les effets de ces forces sont hétérogènes selon le niveau d'intensité technologique des paires de grappes considérées. Le partage d'intrants et de sortants ainsi que le partage de main d'oeuvre contribuent à améliorer la proximité des grappes et donc la colocalisation. Lorsqu'on considère les différentes paires d'intensité technologiques, on constate que l'amplitude du partage d'intrants et de sortants ainsi que du partage de main d'oeuvre est plus faible en valeur absolue pour les paires de grappes dont les deux sont de faible intensité technologique ou de moyenne intensité technologique. Ces résultats coïncident avec l'existant en matière de coagglomération des industries. La relation de citation de brevets ne favorise pas la colocalisation des grappes ce qui est un résultat contraire à celui obtenu avec la coagglomération des industries et laisse alors penser qu'une fois constituées en grappes, les grappes bénéficient des citations de brevets des grappes qui ne sont pas nécessairement situées à proximité.

ANNEXE A

INFORMATIONS ADDITIONNELLES SUR LES DONNÉES

Les matrices de similarités

Contrairement aux autres matrices IO, KS, OES et WFC directement obtenues des bases de données ayant servi à leur construction, des traitements ont été faits sur la base utilisée pour la matrice LM. L'idée est à partir de ces données d'enquêtes de 2000 à 2006, de retenir pour chaque paire d'années, les individus ayant été interviewés la première fois dans l'année la plus ancienne et interviewés également dans l'année la plus récente. La jointure de tels individus se fait sur la base de variables telles que l'identifiant du ménage, le numéro de cet adulte dans le ménage, sa position dans le ménage, le mois de l'entretien, son âge, sa race, son sexe, pour se rassurer d'avoir fait les combinaisons sur le même individu. La variable âge utilisée à cet effet est modifiée en sachant que pour observer le même individu, les autres variables doivent être constantes et la différence d'âge égale 1 entre deux enquêtes. Dans ces bases, des informations comme le code industrie de l'emploi occupé par cet individu entre les deux années sont aussi renseignées. Un empilant les 6 bases obtenues pour chacune des paires successives d'années, on peut alors avoir pour chaque paire de code industrie le nombre d'individus qui entre deux années successives quittent l'emploi dans l'un des codes industries pour l'autre code industrie. Cette valeur permet alors de calculer les coefficients de la matrice LM.

Présentation des données et des codes SCIAN à 4-digits

TABLE A.1: Codes SCIAN 4-digits en 2012

Codes	Description
3111	Fabrication d'aliments pour animaux
3112	Mouture de grains céréaliers et de graines oléagineuses
3113	Fabrication de sucre et de confiseries
3114	Mise en conserve de fruits et de légumes et fabrication de spécialité alimentaires
3115	Fabrication de produits laitiers
3116	Fabrication de produits de viande
3117	Préparation et conditionnement de poissons et de fruits de mer
3118	Boulangeries et fabrication de tortillas
3119	Fabrication d'autres aliments
3121	Fabrication de boissons
3122	Fabrication du tabac
3131	Usines de fibres, de filés et de fils
3132	Usines de tissus
3133	Finissage de textiles et de tissus et revêtement de tissus
3141	Usines de textiles domestiques
3149	Usines d'autres produits textiles
3151	Usines de tricotage de vêtements
3152	Fabrication de vêtements coupés-cousus
3159	Fabrication d'accessoires vestimentaires et d'autres vêtements
3161	Tannage et finissage du cuir et des peaux
3162	Fabrication de chaussures
3169	Fabrication d'autres produits en cuir et produits analogues
3211	Scieries et préservation du bois
3212	Fabrication de placages, de contreplaqués et de produits en bois reconstitués
3219	Fabrication d'autres produits en bois
3221	Usines de pâte à papier, de papier et de carton
3222	Fabrication de produits en papier transformés
3231	Impression et activités connexes de soutien

TABLE A.1 – Codes SCIAN 4-digits en 2012 : suite

Codes	Description
3241	Fabrication de produits du pétrole et du charbon
3251	Fabrication de produits chimiques de base
3252	Fabrication de résines de caoutchouc synthétique et de fibres et de filaments artificiels et synthétiques
3253	Fabrication de pesticides, d'engrais et d'autres produits chimiques agricoles
3254	Fabrication de produits pharmaceutiques et de médicaments
3255	Fabrication de peintures, de revêtements et d'adhésifs
3256	Fabrication de savons, de détachants et de produits de toilette
3259	Fabrication d'autres produits chimiques
3261	Fabrication de produits en plastique
3262	Fabrication de produits en caoutchouc
3271	Fabrication de produits en argile et produits réfractaires
3272	Fabrication de verre et de produits en verre
3273	Fabrication de ciment et de produits en béton
3274	Fabrication de chaux et de produits en gypse
3279	Fabrication d'autres produits minéraux non métalliques
3311	Sidérurgie
3312	Fabrication de produits en acier à partir d'acier acheté
3313	Production et transformation d'alumine et d'aluminium
3314	Production et transformation de métaux non ferreux (sauf l'aluminium)
3315	Fonderies
3321	Forgeage et estampage
3322	Fabrication de coutellerie et d'outils à main
3323	Fabrication de produits d'architecture et d'éléments de charpentes métalliques
3324	Fabrication de chaudières de réservoirs et de contenants d'expédition
3325	Fabrication d'articles de quincaillerie
3326	Fabrication de ressorts et de produits en fil métalliques
3327	Ateliers d'usinage, fabrication de produits tournés de vis, d'écrous et de boulons
3328	Revêtements gravure, traitement thermique et par le froid, et activités analogues
3329	Fabrication d'autres produits métalliques
3331	Fabrication de machines pour l'agriculture, la construction et l'extraction minière

TABLE A.1 – Codes SCIAN 4-digits en 2012 : suite

Codes	Description
3332	Fabrication de machines industrielles
3333	Fabrication de machines pour le commerce et les industries de services
3334	Fabrication d'appareils de ventilation, de chauffage, de climatisation et de réfrigération commerciale
3335	Fabrication de machines-outils pour le travail du métal
3336	Fabrication de moteurs, de turbines et de matériel de transmission de puissance
3339	Fabrication d'autres machines d'usage général
3341	Fabrication de matériel informatique et périphérique
3342	Fabrication de matériel de communication
3343	Fabrication de matériel audio et vidéo
3344	Fabrication de semi-conducteurs et d'autres composants électroniques
3345	Fabrication d'instruments de navigation, de mesure et de commande et d'instruments médicaux
3346	Fabrication et reproduction de supports magnétiques et optiques
3351	Fabrication de matériel électrique d'éclairage
3352	Fabrication d'appareils ménagers
3353	Fabrication de matériel électrique
3359	Fabrication d'autres types de matériel et de composants électriques
3361	Fabrication de véhicules automobiles
3362	Fabrication de carrosseries et de remorques de véhicules automobiles
3363	Fabrication de pièces pour véhicules automobiles
3364	Fabrication de produits aérospatiaux et de leurs pièces
3365	Fabrication de matériel ferroviaire roulant
3366	Construction de navires et d'embarcations
3369	Fabrication d'autres types de matériel de transport
3371	Fabrication de meubles de maison et d'établissement institutionnel et d'armoires de cuisine
3372	Fabrication de meubles de bureau (y compris les articles d'ameublement)
3379	Fabrication d'autres produits connexes aux meubles
3391	Fabrication de fournitures et de matériel médicaux
3399	Autres activités diverses de fabrication

TABLE A.2 – Présentations des sources de données

Données	Disponibilité	Périodicité	Informations disponibles	Remarques
Données Scott's sur toutes les entreprises	2001 à 2019	2 ans	Noms de société, identifiant Scott's, ville, division de recensement, Province, code postal, nom de rue, indicatif régional, téléphone, indicateur Web, Indicateur d'e-mail, nombre d'employés, pieds carrés, estimation Ventes, codes SCIAN, codes SIC, produits, types d'entreprises, nom de marque, année de création, enregistrement ISO, zone métropolitaine, Indicateur d'exportation, indicateur de changement, indicateur du siège social ...	L'année 2015 n'est pas disponible
Matrices canadiennes inputs-outputs	1997 à 2010	1 an	Ce sont des matrices carrées dont les lignes et les colonnes sont Code SCIAN à 6 chiffres	Obtenues des tables inputs-outputs canadiennes
Matrice de citations de brevets			C'est une matrice carrée dont les lignes et les colonnes sont les codes SCIAN à 4-digits	Cette matrice est obtenue de la base de données sur les citations du bureau national pour la recherche économique
Données d'enquêtes courantes sur la population	2001 à 2006	1 an	Cette enquête fournit pour chaque logement sélectionné de variables comme l'identifiant du ménage, le numéro du ménage, la position du ménage dans l'occupation, le mois de l'interview, l'âge, la race, niveau d'étude le code de l'industrie dans lequel il pratique, le code du type d'emploi tenu	Elles sont obtenues du Bureau sur les Statistiques du travail des Etats-Unis
Données d'enquêtes sur les occupations d'emploi	2001 à 2013	2 ans	Donnent des informations comme le nombre de personnes à l'emploi pour chaque paire de code d'emploi et de code SCIAN à 6-digits	Ces données sont obtenues du bureau sur les statistiques du travail des Etats-Unis

ANNEXE B

GÉOCODAGE

Les données sur les établissements dont nous disposons comprennent un ensemble de variables tels que le code postal. Une manière de géocoder ces établissements en vue d'obtenir les coordonnées géographiques serait d'utiliser les fichiers de conversions des codes postaux disponibles chez Statistiques Canada. Ce sont des tables des passages entre les codes postaux et les coordonnées latitudes longitudes. Cependant, le géocodage avec ces tables de passages pose un problème de précision. En effet, surtout en zone rurale, l'espace géographique couvert par un code postal est souvent grand faisant ainsi que deux établissements puissent être associé à un même code postal alors qu'elles sont géographiquement situées à des coordonnées différentes. Aussi, une exploration visuelle des données sur les établissements permettent de remarquer un mauvais placements des établissements comme certains qui se retrouvent positionnées dans l'eau, ou encore décalées de leur adresse réelle. Ces manquements peuvent alors affecter les résultats sur la construction des grappes ce qui justifie donc l'exercice de géolocalisation effectué et que nous allons brièvement décrire ci-dessous.

Le géocodage consiste à transformer une adresse en une paire de coordonnées latitude, longitude pouvant par conséquent être représenté sur une carte géographique et renvoie une meilleure version de l'adresse enregistrées. Entre 2001 et 2017, nous avons 372 333 adresses pour les établissements qui seront utilisées pour le géocodage. Nous avons réalisé trois géocodages

afin de confronter les résultats pour obtenir un meilleur géocodage. Le premier utilise les interfaces de programmations d'applications développées par google en prenant en input l'adresse de l'établissement, le deuxième utilise toujours les interfaces de programmations d'applications développées par google, mais cette fois ci avec en plus de l'adresse le nom de l'entreprise en input, ce qui permet d'obtenir la localisation actuelle de l'établissement même s'il a changé après l'année de collecte et une troisième qui utilise un API alternatif et une base de données sur les adresses qui contient plus de 15 millions d'adresse au Canada.

Ces trois géocodages fournissent chacun des coordonnées géographiques avec pour chaque adresse géocodée un niveau de précision des coordonnées retournées tel que présenté dans le tableau B.1 ci-dessous

TABLE B.1 – Résultats des trois géocodages

		Non trouvé	Approximatif	Centre géométrique	Plage interpolée	Rooftop	Total
Google1	Effectif	18812	25539	14619	96957	1205340	1361267
	%	1.4	1.9	1.1	7.1	88.5	100
Google 2	Effectif	27741	26873	138672	24233	1143748	1361267
	%	2	2	10.2	1.8	84	100
DMTI	Effectif	285293				1075974	1361267
	%	21				79	100

Approximatif correspond aux adresses géocodées avec la précision du code postal, Centre géométrique pour des adresses géocodées avec la précision de centre de polygone, plage interpolée pour les adresses géocodées par interpolation de deux adresses précises, et rooftop pour les adresses géocodées avec la précision jusqu'à l'adresse de la rue.

Les précisions rooftop et range interpolated étant les meilleures, nous ne retenons que les adresses ayant été obtenues avec ce niveau de précision.

Les niveaux de précision approximatif et centre géométrique étant moins précis, nous n'avons pas conservé ces résultats. La règle de décision pour le choix de l'adresse retournée par les différents codages est de comparer les codes postaux obtenu avec les trois géocodages et le code postal renseigné

dans les données Scott's.

TABLE B.2 – Effectifs selon les critères de choix de géolocalisation

Critères	Effectif	%	Choix de géolocalisation
Toutes les options donnent le même code postal	161 958	43.5	Google2
Scott's, Google 1 et 2 donnent le même code postal	39 183	10.52	Google2
Scott's, DMTI et Google 1 donnent le même code postal	4 123	11.07	Google1
Scott's, DMTI et Google 2 donnent le même code postal	20 518	5.51	Google2
Scott's et DMTI donnent le même code postal	18 473	4.96	DMTI
Scott's et Google 1 donnent le même code postal	10 367	2.78	Google1
Scott's et Google2 donnent le même code postal	24 132	6.48	Google2
DMTI et Google 1 et 2 donnent le même code postal	22 625	6.08	Google2
DMTI et Google 1 donnent le même code postal	11 781	3.16	Google1
DMTI et Google 2 donnent le même code postal	1 898	0.51	Google2
Google 1 et 2 donnent le même code postal	11 295	3.03	Google2
Tous les codes postaux sont différents	8 873	2.38	Google1,2 ou DMTI
Total	372 333	100	

Approximatif correspond aux adresses géocodées avec la précision du code postal, Centre géométrique pour des adresses géocodées avec la précision de centre de polygone, plage interpolée pour les adresses géocodées par interpolation de deux adresses précises, et rooftop pour les adresses géocodées avec la précision jusqu'à l'adresse de la rue.

La pertinence du processus de géolocalisation effectué a été apprécié en calculant le pourcentage d'adresse mieux répertorié en référence aux codes postaux déjà renseigné dans la base Scott's. Ainsi, avec les trois géocodage, 1065060 soit 80% des tous les géocodages ont donné une adresse avec le même code postal que celui répertorié dans les Scott's, 2.7 % soit 36281 donnent des codes postaux avec la même région de tri d'acheminement (les trois premiers chiffres du code postal) que le code postal donné dans les scott's et la distance bilatérale entre les deux coordonnées géographiques inférieures à 250 mètres et 17.3% pour les adresses n'ayant pas pu être améliorée. Ce qui attribue un score de plus de 80% de précision des coordonnées obtenues à l'issues du géocodage.

ANNEXE C

CONCORDANCE SCIAN

TABLE C.1 – Choix du codes de concordance des SCIAN 6-digits entre 1997 et 2017

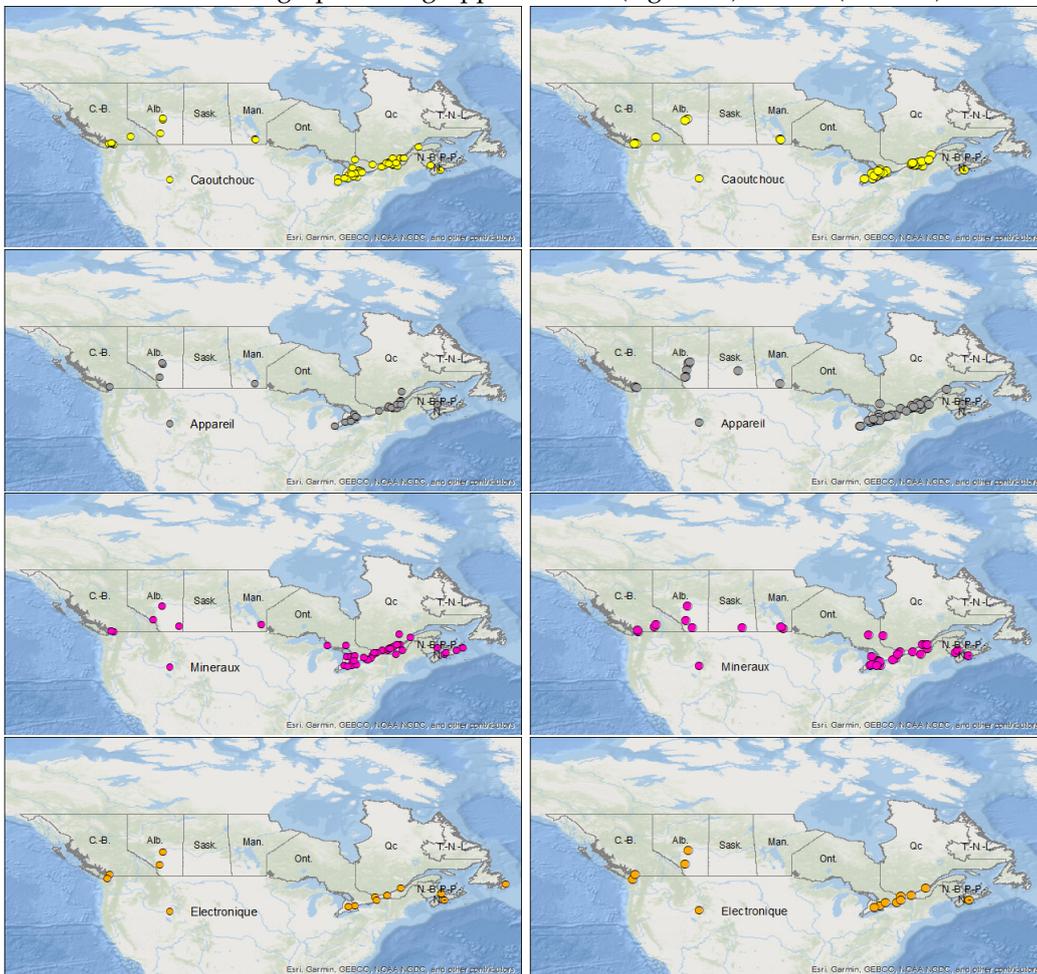
1997	2002	2007	2012	2017	Choix	Années du Choix
326198	326198	326196	326196	326196	326198	1997, 2002
326198	326198	326198	326198	326198	326198	1997, 2002
333299	333299	333299	333248	333248	333299	1997, 2002, 2007
311320	311320	311320	311351	311351	311320	1997, 2002, 2007
311330	311330	311330	311352	311352	311330	1997, 2002, 2007
311822	311822	311822	311824	311824	311824	2012
311823	311823	311823	311824	311824	311824	2012
315221	315221	315221	315220	315220	315220	2012
315222	315222	315222	315220	315220	315220	2012
315226	315226	315226	315220	315220	315220	2012
315227	315227	315227	315220	315220	315220	2012
315229	315229	315229	315220	315220	315220	2012
315231	315231	315231	315249	315249	315249	2012
315232	315232	315232	315249	315249	315249	2012
315233	315233	315233	315249	315249	315249	2012
315234	315234	315234	315249	315249	315249	2012
315239	315239	315239	315249	315249	315249	2012
315291	315291	315291	315241	315241	315291	1997, 2002, 2007
315292	315292	315292	315281	315281	315292	1997, 2002, 2007
315299	315299	315299	315289	315289	315299	1997, 2002, 2007
333210	333210	333210	333245	333245	333210	1997, 2002, 2007
333220	333220	333220	333246	333246	333220	1997, 2002, 2007
333291	333291	333291	333247	333247	333291	1997, 2002, 2007
337123	337123	337123	321999	321999	321999	2012

Ces codes à 4-digits sont assez stables entre 1997 et 2017

ANNEXE D

AUTRES STATISTIQUES ET RÉSULTATS

FIGURE D.1 – Cartographie des grappes en 2001 (à gauche) et 2019 (à droite) : suite



Source : Calculs de l'auteur

TABLE D.1 – Emploi manufacturier par année et province au Canada

		Provinces										
		NFL	PEI	NS	NB	QC	ON	MAN	SASK	AB	BC	Canada
2001	N	19.2	5.6	41.4	41.2	554.4	922.5	65.7	26.4	130	170.8	1977.6
	%	0.97	0.28	2.09	2.08	28.03	46.64	3.32	1.33	6.58	8.64	100
2003	N	17.10	5.1	39.4	38.6	539.6	896.3	64.9	25.6	122.1	162.3	1911.4
	%	0.89	0.26	2.06	2.02	28.23	46.89	3.39	1.34	6.39	8.49	100
2005	N	14.5	6.1	38.2	33.8	511.1	854.4	62.6	27.6	128.4	163.8	1840.9
	%	0.79	0.33	2.07	1.83	27.76	46.41	3.4	1.5	6.98	8.9	100
2007	N	15.6	x	33.5	32.4	484.4	800	62.6	26.6	138	165.7	1764.6
	%	0.88	x	1.9	1.84	27.45	45.34	3.55	1.51	7.82	9.39	100
2009	N	11.2	5	31.3	28.7	424.5	654.9	55.3	25.1	118.7	138.2	1493.3
	%	0.75	0.34	2.1	1.92	28.43	43.86	3.7	1.68	7.95	9.25	100
2011	N	11.4	5.1	31.2	30.5	419.2	654.7	54.4	24.7	129	138.6	1499.3
	%	0.76	0.34	2.08	2.04	27.96	43.67	3.63	1.65	8.6	9.24	100
2013	N	11.2	4.8	30.4	29.2	408.4	660.6	53.3	28.2	136.1	140	1502.6
	%	0.74	0.32	2.02	1.94	27.18	43.96	3.54	1.88	9.06	9.32	100
2017	N	10.2	5.1	30.6	28.9	418.6	670.2	53.2	23.8	119.2	149.4	1509.6
	%	0.67	0.34	2.03	1.91	27.73	44.39	3.52	1.58	7.9	9.89	100
2019	N	9.8	5.4	31.7	31.4	447.8	692.4	57	25.2	123.9	155.6	1580.7
	%	0.62	0.34	2.01	1.99	28.33	43.8	3.61	1.59	7.84	9.85	100

Source : <https://www150.statcan.gc.ca/t1/tbl1/fr/tv.action?pid=1410020201>.
N est l'emploi en millier et arrondi au millier le plus proche. x : donnée indisponible du fait de la confidentialité

TABLE D.2 – Emploi et nombre d'entreprises par groupes d'industries et par années

year	2001	2003	2005	2007	2009	2011	2013	2017	2019
Aliments	259.9	268	270.9	276.8	264.6	264.6	243.1	184.7	166.2
	5.1	5	4.7	4.4	4.3	4.3	3.6	2.9	2.6
Appareil	231.2	160.7	163.9	164.4	160.3	164	151.1	190.8	105.1
	4.5	4.6	4.8	4.7	4.5	4.3	3.7	4.4	2.8
Bois	246.1	255.6	260.8	238.4	232.4	201	180.8	116.4	118.5
	7.1	7.1	6.6	6	5.7	5.3	4.2	2.8	2.8
Caoutchouc	120	128.9	141.1	127.1	120.8	108.7	106.3	95.7	83.9
	2.5	2.6	2.5	2.4	2.3	2.2	2	1.9	1.6
Charbon	126.2	131.2	145.9	149.1	152.1	147	133	98.6	92.8
	2.4	2.3	2.2	2.1	2.1	2	1.7	1.4	1.2
Electronique	149.2	125.1	96.8	97.1	99.1	87.1	73.1	57.6	54.3
	2.1	1.8	1.6	1.6	1.6	1.5	1.3	0.9	1
Mineraux	126.1	74.9	70.7	72.8	101.6	67.1	62.8	80.1	57
	2.5	2.6	2.6	2.4	2.4	2.2	1.8	1.8	1.3
Papier	106.8	132	101.9	106.7	87.6	78.7	64.5	48.8	44.1
	0.9	0.9	0.8	0.8	0.8	0.7	0.6	0.5	0.4
Textile	282.9	241.8	226.2	199.4	183	173.6	132.2	81.8	100.1
	9.4	9.2	8.3	7.3	6.9	6.4	5.2	3.6	3.5
Transport	298.5	240.6	238.4	231.5	216.8	202	178.6	143.4	127.4
	1.7	1.7	1.6	1.6	1.5	1.4	1.2	0.9	0.8
Total	1946.9	1758.9	1716.6	1663.2	1618.2	1493.8	1325.4	1097.8	949.3
	38	37.7	35.8	33.3	32	30.2	25.3	21.2	18.1

Source : Calculs de l'auteur, Scott's

TABLE D.3 – Emploi manufacturier par année et secteur au Canada

NAICS	2001	2003	2005	2007	2009	2011	2013	2017	2019	μ_s	$^*\mu_c$	d
11	13.7	29.3	32.8	24.5	25.4	24.2	21.3	16.5	14.6	22.5	48.8	-53.9
21	4.1	23.2	52.4	82.1	93.9	107.7	134	94.3	99.5	76.8	184.2	-58.3
22	60.1	73.7	46.3	40.1	46	62.8	92.1	80.6	60.6	62.5	117.7	-46.9
23	73.1	66.6	83.5	158.4	169.7	197.8	284.4	288.3	262.2	176	814.4	-78.40
31-33	2495.30	2280.70	2174.1	2116.4	2062.1	1897	1726.8	1426	1243.90	1935.8	1675.6	15.5
41-45	753.2	619.9	669.1	851.5	873.9	937.9	928.6	876.3	687.9	799.8	2579.20	-69
48-49	105.8	138.20	179.2	240.8	280.60	315.60	359.4	325.60	315.90	251.2	672.6	-62.7
51	284	225.1	242	267.2	285.2	327	366.7	272.7	252.6	280.3	336	-16.60
52	245.3	163.1	183.1	229.3	281	361.7	401.5	410.3	350.8	291.8	665.2	-56.1
53	44.8	47.3	60.6	84.2	89.1	114.1	150.20	137.20	155.4	98.1	261.90	-62.5
54	237.6	186.8	245.2	465.4	551.1	637.9	730.8	868.1	806	525.5	770.4	-31.8
55	2.5	3.4	11.7	15.6	42.7	44.7	7.5	2.8	2.20	14.8	102.8	-85.6
56	49	56.5	86.8	117.2	126.5	196.8	228.7	214.9	177.9	139.4	722.9	-80.7
61	34.70	42.5	44.7	1079.10	110.1	150.30	185.2	174.5	250.9	230.2	1147.7	-79.90
62	235.8	299.7	382.1	303.2	331.3	193.3	30	357.4	0.4	237	1623.3	-85.4
71	17.10	18.10	19.10	34.9	37.9	73.7	78.5	87.1	72.40	48.8	250.8	-80.5
72	37.9	38.70	51.3	69.6	71.40	103.9	133.1	145.20	128.30	86.6	1103.7	-92.2
81	70.5	22.4	46.8	108.5	149.70	196.7	261.40	279.3	193	147.6	507.4	-70.90
91	0	0	0	0	0	0	41.8	65	68.8	58.5	1001.8	-94.2

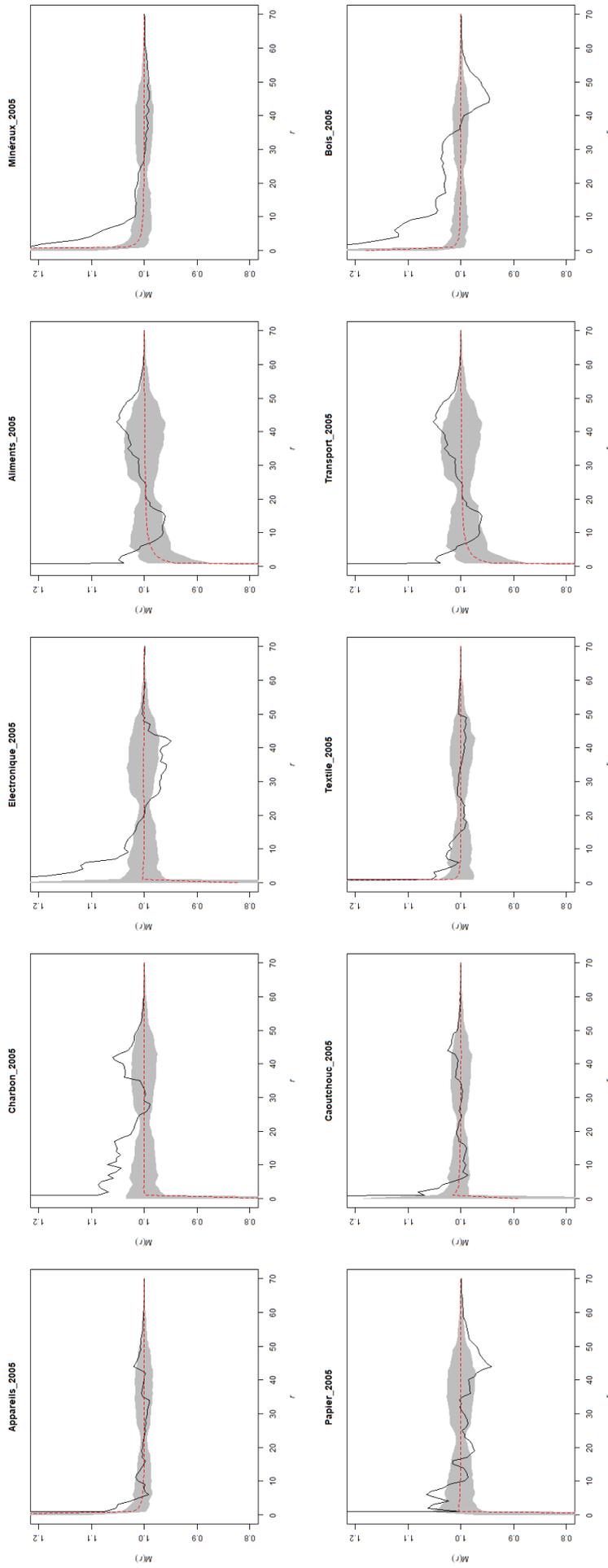
Source : Calculs à partir de la base de données Scott's. L'emploi est en millier et arrondi au millier le plus proche. μ_s respectivement μ_c représentent la moyenne sur données scott's et sur données de statistique Canada (Tableau D.4)

TABLE D.4 – Emploi manufacturier par année et secteur au Canada

SCIAN	2001	2003	2005	2007	2009	2011	2013	2017	2019
11	68.1	63	59.1	52.5	39.5	40.2	38.3	39.6	38.5
21	138.9	149	162.5	192.2	181.6	207.1	226	199.8	200.5
22	107.2	108.7	116.4	118.2	118.6	123.5	118.9	120.2	127.5
23	577.7	631.3	692.1	783.6	798	861.4	947.1	997.3	1041.2
31-33	1977.6	1911.4	1840.9	1764.6	1493.3	1499.3	1502.6	1509.6	1580.7
41-45	2261.9	2395.5	2458.3	2605.3	2610	2615.5	2681.1	2753.8	2831.6
48-49	590.2	606.3	631.9	665.9	660.3	673.4	705.5	738.2	781.4
51	327.8	320.5	321.7	336.3	333.2	343.7	342.9	352.4	345.9
52	564	588.8	637.8	670.7	685.5	699	725.6	749.8	
53	228.3	244.2	255.4	251.9	254.3	269.3	294.4	297.3	
54	643.8	645.5	672.4	743.2	756.3	786	821.8	891.1	973.8
55	93.1	95	98.3	107.7	112.8	104.2	107.1	97.5	109.4
56	557.2	624.2	699.7	762.9	723.6	743.1	764.9	801	829.8
61	981.3	1022.8	1061	1111.4	1154	1159.9	1203.8	1279.4	1355.8
62	1318.4	1383.3	1426.8	1500.6	1604.3	1681.7	1746.3	1917.7	2030.9
71	205.8	223.1	230.6	242.2	249.3	248	252.4	293.5	312.3
72	944.6	971.9	977	1045.8	1074.4	1099.2	1178.7	1299.9	1342.2
81	455.8	470.2	478.1	498.4	507.7	521	537.9	541	556.6
91	857.2	896.5	926.1	967.4	1034.6	1057	1041	1089.3	1147.2

Source : <https://www150.statcan.gc.ca/t1/tb11/fr/tv.action?pid=1410020201>. L'emploi est en millier et arrondi au millier le plus proche.

FIGURE D.2 – Fonction M pour chaque groupes d'industries en 2005, r en km



Source :

BIBLIOGRAPHIE

- Baptista, R. & Swann, P. (1998). 'Do firms in clusters innovate more?'. *Research Policy* 27(5) :525–540.
- Behrens, K. (2016). 'Agglomeration and clusters : Tools and insights from coagglomeration patterns'. *Revue canadienne d'Économique* 49(4) :1293-1339.
- Behrens, K. Boualam, B. & Martin, J. (2020). 'Are clusters resilient ? Evidence from Canadian textile industries'. *Journal of Economic Geography* 20(1) :1–36.
- Behrens, K. & Bougna, T. (2015). 'An anatomy of the geographical concentration of Canadian manufacturing industries'. *Regional Science and Urban Economics* 51,47-69.
- Chinitz, B. (1961). 'Contrasts in Agglomeration : New York and Pittsburgh'. *The American Economic Review* 51(2) :279–289.
- Delgado, M. Porter, M. E. & Stern, S. (2010). 'Clusters and entrepreneurship'. *Journal of Economic Geography* 10(4) :495–518.
- Delgado, M. Porter, M. E. & Stern, S. (2016). 'Defining clusters of related industries'. *Journal of Economic Geography* 16(1) : 1–38.
- Dongen, S. (2000). A cluster algorithm for graphs. Technical Report. Centre for Mathematics and Computer Science.

- Duranton, G. (2010). *The economics of clusters : lessons from the French experience*. Studies of policy reform, University Press, Oxford.
- Duranton, G. & Overman, H. G. (2005). 'Testing for Localization Using Micro Geographic Data'. *Review of Economic Studies* 72(4) :1077–1106.
- Ellison, G. & Glaeser, E. L. (1997). 'Geographic Concentration in U.S. Manufacturing Industries : A Dartboard Approach'. *Journal of Political Economy* 105(5) : 889–927.
- Ellison, G. Glaeser, E. L. & Kerr, W. R. (2010). 'What Causes Industry Agglomeration ? Evidence from Coagglomeration Patterns'. *The American Economic Review* 100(3) :1195–1213.
- EOCIC (2019).
URL: <https://www.eucluster2019.eu/files/events/4538/files/eocic-cluster-programme-report-2905.pdf>
- Everitt, B. S. Landau, S. Leese, M. & Stahl, D. (2011). *Cluster Analysis*. John Wiley & Sons.
- Faggio, G. Silva, O. & Strange, W. (2017). 'Heterogeneous Agglomeration'. *The Review of Economics and Statistics* 99(1) :80–94.
- Feldman, M. & Audretsch, D. (1999). 'Innovation in cities : Science-based diversity, specialization and localized competition'. *European Economic Review* 43(2) :409–429.
- Gabaix, X. & Ibragimov, R. (2011). Rank 1/2 : A Simple Way to Improve the OLS Estimation of Tail Exponents'. *Journal of Business & Economic Statistics* 29(1) :24–39.
- Gabaix, X. & Ioannides, Y. M. (2004). Chapter 53 - The Evolution of City Size Distributions. in J. V. Henderson & J.-F. Thisse, eds, 'Handbook of Regional and Urban Economics'. Vol. 4 of *Cities and Geography* :2341–2378.

- Glaeser, E. L. (1999). 'Learning in Cities'. *Journal of Urban Economics* 46(2) :254–277.
- Hatzichronoglou, T. (1997). 'Révision des classifications des secteurs et des produits de haute technologie'. OCDE.
- Helsley, R. W. & Strange, W. C. (1990). 'Matching and agglomeration economies in a system of cities'. *Regional Science and Urban Economics* 20(2) :189–212.
- Howard, E., Newman, C. & Tarp, F. (2016). 'Measuring industry coagglomeration and identifying the driving forces'. *Journal of Economic Geography* 16(5) :1055–1078.
- Jacobs, J. & Jane, J. (1969). *The Economy of Cities*.
- Krugman, P. (1991). 'Increasing Returns and Economic Geography'. *The Journal of Political Economy* 99(3) : 483–499.
- Marcon, E. & Puech, F. (2010). 'Measures of the geographic concentration of industries : improving distance-based methods'. *Journal of Economic Geography* 10(5) :745–762.
- Marshall, A. (1920). 'Principles of Economics (8th ed.) - Online Library of Liberty'.
- Martin, P. Mayer, T. & Mayneris, F. (2016). Are clusters more resilient in crises ? Evidence from French exporters in 2008-2009. Sciences Po publications.
- Martin, R. & Sunley, P. (2003). 'Deconstructing clusters : chaotic concept or policy panacea ?'. *Journal of Economic Geography* 3(1) :5–35.
- Porter, M. (1990). 'The Competitive Advantage of Nations'. *Harvard Business Review* 68(2) :73–93.

Roy, J. (2010). 'Impact des coûts de transport sur la compétitivité des entreprises québécoises'. p. 38.

Vernon, R. (1960). *Metropolis 1985 : an Interpretation of the Findings of the New York Metropolitan Region Study*. Harvard University Press.