

L'interdépendance spatiale de l'investissement américain dans les économies provinciales canadiennes

par

Denis Poulin

mémoire présenté à la
Faculté des études supérieures et postdoctorales
pour l'obtention
du grade de Maître ès arts
en Science économique

Département de science économique
Faculté des Sciences sociales
Université d'Ottawa

© Denis Poulin, Ottawa, Canada, 2013

Résumé

L'étude empirique de l'investissement direct étranger à l'aide des modèles économétrique spatiale est maintenant bien établie. Cependant, peu d'études se sont penchées sur l'investissement direct étranger dans des économies développées et un nombre encore plus restreint traite du cas où ces économies sont celles de provinces d'un même pays. Ce mémoire tente de pallier cette lacune en étudiant l'investissement direct américain entre 2004 et 2011 dans les provinces canadiennes et en utilisant le modèle spatial SAC qui combine le modèle autorégressif spatial et le modèle d'erreur spatiale. Nos résultats démontrent une grande robustesse de la présence de l'interdépendance spatiale de l'investissement direct américain. De plus, ils proposent que l'argument de compétitivité utilisé pour justifier la baisse du niveau de taxation des entreprises est sans fondement dans le cas des provinces canadiennes.

Avant-propos

Je veux, tout premièrement, remercier ma conjointe, mon amour de toujours, Myriam, pour son soutien inconditionnel et sa patience quasi divine. Je remercie également mes enfants qui malgré leurs jeunes âges ont bien voulu accepter que leur père fasse des devoirs avant de pouvoir passer du temps avec eux.

Ce mémoire n'aurait pu voir le jour sans l'aide et l'appui de Jean-François Tremblay, mon directeur de mémoire. Ce dernier s'est avéré un mentor en économie que je qualifie de génial. Son temps, sa patience et ses conseils ont été plus qu'utiles et appréciés.

Table des matières

Résumé	i
Avant-Propos	iii
Table des matières	vi
Liste des figures	vii
Liste des tableaux	x
1 Introduction	1
2 Revue de la littérature	5
2.1 Tobin- Q	7
2.1.1 La théorie	7
2.1.2 Extension : L'incertitude et l'irréversibilité	12
2.1.3 Problème dynamique des données	13
2.2 Théorie de la gravité	15
2.2.1 L'intuition derrière la théorie de la gravité et ses grandes lignes	15
2.2.2 La présence de 0	17
2.2.3 Le modèle de gravité et la taxation corporative	20
2.3 Économétrie spatiale	22
2.3.1 Problème généraux de l'économétrie spatial	22
2.3.2 Modèle SAR	25
2.3.3 Modèle d'erreur spatiale	34
2.3.4 Panel non-balancé	36

3	Données	39
3.1	L'investissement direct américain au Canada	39
3.2	Description	41
4	Spécification des modèles	47
4.1	Modèle de gravité standard	47
4.2	Spécifications spatiales	49
5	Résultats empiriques	53
5.1	Sélection du modèle	53
5.2	Modèle d'investissement et variables de contrôle	57
5.3	Marché visé	61
6	Conclusion	65
	Bibliographie	67
A	Liste des états américains limitrophes du Canada	73
B	Tableaux de résultats des régressions	75
C	Tableaux descriptifs supplémentaires	81
D	Figures supplémentaire	85

Table des figures

2.1	Fonction ψ	9
5.1	Répartition de l'investissement direct américain dans les provinces canadiennes en 2011	59
D.1	Distribution des résidus du modèle de base pour la province de l'Alberta	85
D.2	Distribution des résidus du modèle de base pour la province de la Colombie Britannique	86
D.3	Distribution des résidus du modèle de base pour la province du Manitoba	86
D.4	Distribution des résidus du modèle de base pour la province du Nouveau-Brunswick	87
D.5	Distribution des résidus du modèle de base pour la province de la Nouvelle-Écosse	87
D.6	Distribution des résidus du modèle de base pour la province de l'Ontario	88
D.7	Distribution des résidus du modèle de base pour la province du Québec	88
D.8	Distribution des résidus du modèle de base pour la province de la Saskatchewan	89
D.9	Distribution des résidus du modèle de base pour la province de Terre-Neuve-Labrador	89

Liste des tableaux

2.1	Résultats de l'estimation de l'équation de base de [9]	20
2.2	Signe des différents paramètres du modèle SAR d'investissement selon le modèle d'investissement	30
3.1	Moyenne annuelle du pourcentage de l'investissement direct américain de l'investissement direct étranger de 2004 à 2011	40
3.2	Statistiques descriptives	44
5.1	Test d'autocorrélation spatiale du multiplicateur de Lagrange Baltagi et col. [8]	55
B.1	Analyse spatiale de l'investissement direct américain - Sélection du modèle .	76
B.2	Analyse spatiale de l'investissement direct américain - Variables de contrôles excluant METR	77
B.3	Analyse spatiale de l'investissement direct américain - Variables de contrôles avec METR	78
B.4	Analyse spatiale de l'investissement direct américain - Marché visé excluant METR	79
B.5	Analyse spatiale de l'investissement direct américain - Marché visé avec METR	80
C.1	Statistiques descriptives pour la province de l'Alberta	81
C.2	Statistiques descriptives pour la province de la Colombie Britannique	82
C.3	Statistiques descriptives pour la province du Manitoba	82
C.4	Statistiques descriptives pour la province du Nouveau-Brunswick	82
C.5	Statistiques descriptives pour la province de la Nouvelle-Écosse	83

C.6	Statistiques descriptives pour la province de l'Ontario	83
C.7	Statistiques descriptives pour la province du Québec	83
C.8	Statistiques descriptives pour la province de la Saskatchewan	84
C.9	Statistiques descriptives pour la province de Terre-Neuve-Labrador	84

Chapitre 1

Introduction

L'importance de l'investissement direct étranger au Canada n'est plus à discuter comme le montre sa progression de plus de 27% entre 2000 et 2011. L'essor exponentiel du secteur des ressources naturelles au Canada s'est concrétisé durant cette période donnant ainsi lieu aux investissements de plusieurs multinationales étrangères. Des modèles économiques expliquant le comportement de ces multinationales ont fait leurs apparitions en 1984 avec Markusen [58], modèle horizontal, et Helpman [39], modèle vertical. Tous deux proposent des modèles d'équilibre général comprenant des investissements bilatéraux dans le cadre de la théorie de la gravité. Ce n'est qu'avec l'article révolutionnaire de Coughlin et Segev [29] que l'importance de la situation géographique de la localisation des investissements est abordée et étudiée empiriquement dans le cadre de l'investissement direct étranger. Par la suite, d'autres types de modèle d'investissement voient le jour comme celui de la plateforme d'exportation [33], [73], [12] et le modèle vertical complexe [6]. Ces modèles développés dans le cadre de la théorie de la gravité mettent en lumière un enjeu important : les marchés voisins de la localisation de l'investissement, communément appelé "l'effet du troisième pays". Cette nouvelle approche implique une interdépendance spatiale de l'investissement direct étranger. Ainsi l'investissement localisé dans un pays dépend également de l'investissement fait chez ses voisins avec une influence diminuant avec la distance séparant les pays. L'importance de l'effet du troisième pays n'est presque plus à discuter depuis la parution de [15] et qui a stimulé la recherche empirique sur ce sujet.

La littérature empirique considérant l'interdépendance spatiale comme un facteur non négligeable et déterminant de l'investissement direct étranger est de plus en plus volumineuse. Cependant à la connaissance de l'auteur, peu d'études portent sur l'investissement direct étranger dans les provinces ou régions d'un même pays autre que la Chine.¹ Dans ce contexte, l'accessibilité des données est très problématique comme le démontre [65] qui est la première étude disposant de données désagrégées au niveau des industries de Chine, l'un des pays les plus populaires dans la littérature. Ce mémoire se veut donc être un outil afin de palier à la lacune d'études empiriques sur l'investissement direct étranger localisé dans les provinces d'un même pays. Nous profitons de la qualité et de la disponibilité relative des données au Canada pour contourner la deuxième problématique soulevée précédemment. Cette étude en utilisant des modèles empiriques spatiaux souhaite déterminer quelles sont les motivations générales des investissements directs américains dans les provinces canadiennes et plus particulièrement si le marché canadien est le principal marché visé par cet investissement. De par sa situation géographique, le Canada donne un accès d'envergure au marché américain. L'accord de libre-échange entre ces deux pays ne fait qu'augmenter la probabilité que la proximité des États-Unis puisse jouer un rôle important dans la prise de décision de la localisation de l'investissement. L'importance du marché visé est primordiale dans l'élaboration des politiques reliées à l'investissement. Cette étude peut ainsi servir de point de départ à une réflexion portant sur la mise en place ou non de l'élaboration de politiques permettant une meilleure accessibilité au marché canadien et ainsi concurrencé efficacement le marché américain.

Le deuxième objectif visé par cet ouvrage est de connaître si le niveau de taxation des entreprises a un réel impact sur le processus décisionnel de la localisation de l'investissement américain. L'argument avançant que les pays doivent abaisser le taux de taxation des corporations afin de demeurer compétitif² sur le plan international est ici remis en question lorsque ce dernier est adressé à des gouvernements provinciaux. Les enjeux de compétitivité nationaux et internationaux sont très différents. Dans un contexte international, une multina-

1. Les articles suivants sont ceux connus de l'auteur et où les hôtes de l'investissement sont des provinces d'un même pays : [18], [34], [62], [46] et [47].

2. Cet argument est présenté entre autre dans [24] et [27]

tionale doit faire un choix entre des économies développées, émergentes et sous-développées. Avec un coût de main-d'oeuvre et de production moins élevé, les pays émergents sont favorisés. De plus, ces pays offrent une accessibilité à certaines infrastructures et un niveau acceptable de la qualité de la main-d'oeuvre. Les pays développés quant à eux doivent valoriser leurs infrastructures et diminuer les coûts de production des entreprises en abaissant entre autres choses le niveau de taxation afin de demeurer compétitifs. La situation étudiée dans ce mémoire est très différente puisque toutes les provinces canadiennes sont des économies développées et que le choix du pays est déjà effectué. Il est certain qu'un taux bas de taxation est un avantage, cependant la localisation des ressources nécessaires à la production de même que l'accessibilité à des infrastructures publiques peuvent possiblement contrebalancer les effets d'un niveau plus élevé de taxation. Les gouvernements peuvent surenchérir en mettant en place des politiques d'amélioration de la qualité de la main d'oeuvre et assurer ainsi des gains sur le plan de la productivité. Cet enjeu est crucial pour les gouvernements, car si ceux-ci fixent le niveau de taxation sous le seul nécessaire au maintien des infrastructures et de la formation de la main-d'oeuvre, les investissements diminueront au lieu d'augmenter produisant ainsi l'effet contraire de celui escompté.

Dans le but de répondre à ces questions, nous utilisons un ensemble de données portant sur l'investissement direct américain dans les provinces canadiennes, excluant l'Île-du-Prince-Édouard³, entre 2004 et 2011. Nous estimons à l'aide de la méthode du maximum de vraisemblance un modèle spatial SAC, modèle combinant le modèle autorégressif spatial (SAR) et le modèle d'erreur spatial (SEM), et incluons seulement des effets fixes provinciaux. La variation dans nos variables de contrôle nous montre une très grande robustesse de la significativité de l'interdépendance spatiale observée. Ces résultats sont similaires à ceux de [35] et [71]. Nous déterminons que la proximité des fournisseurs et des ressources est corrélée positivement avec l'investissement qui suit ainsi un modèle de type vertical complexe. Cet effet de proximité est renforcé par un effet d'agglomération qui se traduit par des estimés relativement assez élevées des coefficients des variables mesurant le potentiel des marchés voisins. Fait inusité, les chocs aléatoires au plan spatial et temporel modélisés de manière à

3. Voir Section 3.1 pour une explication complète de ce point.

ce qu'ils soient interdépendants spatialement poussent à la baisse les investissements directs américains. Le manque de coordinations dans l'implantation de politiques ayant pour but de stimuler l'économie et les investissements comme ceux mis en place lors de la récession de 2008-2009 est l'une des explications de ce phénomène. Nous avons également observé que la proximité des marchés américains est un facteur non négligeable. Finalement, un test de robustesse indique que nos résultats sont sensibles aux types de mesure de distance. De tels résultats dans le cadre de l'économétrie spatiale sont courants lorsque la structure spatiale est modifiée. Cependant, nous n'avons fait que varier la mesure et non la structure spatiale.

Le mémoire se poursuit de la manière suivante. Le chapitre 2 présente une revue de la littérature qui s'articule en deux temps. Nous faisons un survol des principales théories microéconomiques reliées à l'investissement. Nous en faisons par la suite une critique en utilisant des exemples concrets d'études empiriques. La dernière section de ce chapitre traite des différents modèles économétriques spatiaux qui ont vu leurs popularités croître de manière exponentielle durant la dernière décennie. Le Chapitre 3 donne une description des données que nous utilisons et des constructions que nous avons effectuées. Nous poursuivons au chapitre 4 avec la présentation du modèle économétrique SAC que nous avons choisi ainsi que des différentes variantes que nous avons estimées. Le chapitre 5 se compose premièrement des tests et estimés relatifs au choix du modèle économétrique. Nous abordons ensuite les résultats des différentes spécifications reliées à l'étude des problématiques précédemment mentionnées. Une conclusion termine ce mémoire.

Chapitre 2

Revue de la littérature

Il est indéniable que la mondialisation des dernières décennies a fondamentalement changé le portrait de l'économie mondiale. La part de l'investissement des multinationales dans les économies émergentes s'est accrue de manière exponentielle. L'un des exemples frappants de ce phénomène est la Chine qui avec sa politique "de portes ouvertes" adoptée en 1970 à propulser la Chine au premier rang des pays hôte d'investissement direct étranger passant de 0,63 milliard en 1983 à près de 116,0 milliards (dollar américain).¹

Étant vue comme un des moteurs de croissance économique, l'étude de l'investissement direct étranger a motivé plusieurs modèles économiques théoriques. De plus avec une plus grande accessibilité aux données et une technologie plus avancée, les études empiriques de ces théories de l'investissement se sont multipliées au cours des 20 dernières années. Le but de ce chapitre est en un premier temps d'exposer les théories microéconomiques les plus couvertes durant les dernières décennies et d'en faire la critique par des exemples concrets d'études empiriques. Avec l'avènement des ordinateurs performants, l'économétrie a connu une croissance impressionnante et l'investissement est l'un des domaines où les économétriciens sont les plus actifs. Nous traitons donc des modèles économétriques spatiaux qui touchent à l'investissement direct étranger. La dimension spatiale n'est apparue que très récemment dans l'étude de l'investissement, mais est maintenant de plus en plus répandue!

Ce chapitre se poursuit comme suit. Nous amorçons avec la section 2.1 notre étude de la théorie de l'investissement par exposer la Théorie- q introduite par Tobin [67]. Pour ce faire,

1. Données du Bureau National de la Statistique de Chine.

nous préconisons l'approche de la maximisation d'accumulation de capital tenant compte du facteur d'ajustement des firmes. Une extension de ce modèle est présentée par la suite tenant compte de l'irréversibilité de l'investissement et de l'incertitude. Ce chapitre se termine avec une critique de ce modèle du point de vue empirique en étudiant l'article de Blunbel et coll. [17].

La section 2.2 traite du modèle de gravité. Nous discutons de son importance et des motivations de l'introduction de la dimension de distance qui est au coeur de ce modèle. La sous section 2.2.2 aborde le problème de la présence de 0 dans les données et des possibilités de résoudre ce problème de façon empirique avec l'étude des articles [47] et [70]. Par la suite nous présentons comment le modèle de gravité peut être utilisé pour étudier les impacts d'une taxation corporative. Nous utilisons l'article de Benassy-Queré et coll. [9] qui traite de façon remarquable comment l'utilisation de différentes mesures du taux de taxation affecte les estimations.

Le dernière section se penche sur des modèles plus récents fréquemment utilisés dans la littérature empirique. Nous commençons par traiter de la dimension spatiale et des impacts de son introduction en économétrie. Deux problèmes sont abordés soit l'hétérogénéité spatiale et la dépendance spatiale. La sous section 2.3.2 étudie en détail le modèle spatial autorégressif. Une exposition du modèle est faite en un premier temps et est suivie par une interprétation des signes de différents paramètres à estimer et des modèles d'investissement de firmes multinationales. De par sa spécification, ce modèle nécessite des modifications aux méthodes d'estimation classique. Nous étudions donc ce point en mettant un accent particulier sur des articles utilisant des données de panel qui ont eu des retombées importantes sur l'étude de l'effet du 3^e pays. Cette partie se termine par l'étude de la thèse de doctorat de Wei [71] qui propose une technique d'estimation sur des panels non balancés. Cette section traite finalement d'un deuxième modèle en économétrie spatiale, le modèle d'erreur spatiale. Étant donné le peu de place qu'occupe ce modèle dans la littérature empirique, nous traitons ce modèle que par l'étude d'un seul article [32] qui lui aussi propose une technique d'estimation pour des données de panel spatial non balancé.

2.1 Tobin- Q

Entre 1970 et 1980, deux théories économiques ont dominé la littérature portant sur l'investissement. La première originalement proposée par Jorgenson [42] se base sur la théorie néoclassique. La deuxième se nomme Théorie- q et fut introduite par Tobin [67]. Cette section traite exclusivement de la seconde théorie. En un premier temps, nous exposons la Théorie- q , sous la perspective néoclassique [38]. Par la suite, nous poursuivons avec une revue de certains articles empiriques de la dernière décennie pour examiner comment la Théorie- q peut être adaptée et testée empiriquement par rapport à

1. l'incertitude de l'investissement incorporant les coûts fixes d'implantation de l'investissement et de potentiels investissements irréversibles ;
2. l'interaction de la dynamique des mesures quantitatives et les impacts de la taxation sur l'investissement ;
3. l'imperfection du marché du capital.

2.1.1 La théorie

Dans [67], Tobin suggère que le taux d'investissement est une fonction du ratio entre la valeur du marché de l'ajout du capital d'un bien d'investissement et de son coût de remplacement. Tobin nomme ce ratio q parfois nommé *shadow price* dans la littérature. Il est certain que les firmes font ainsi face à un coût d'ajustement, point présent dans l'approche néoclassique. Elles augmentent ou diminuent leurs stocks de capital de façon à ce que q aie une valeur de 1. Il est maintenant connu que la Théorie- q est équivalente au problème d'optimisation d'accumulation de capital des firmes en considérant bien entendu le coût d'ajustement. Nous présentons donc la Théorie- q en utilisant la perspective néoclassique.² Nous suivons de très près Hayashi qui, dans [38] traite cette équivalence très rigoureusement.

Soit $r(s)$ le taux d'escompte nominal. Soit $\pi(t)$ le profit de l'entreprise au temps t donné

2. Celle-ci est parfois nommée dans la littérature comme étant la théorie néoclassique modifiée de l'investissement.

par

$$\pi(t) = p(t)F(K(t), L(t))w(t)L(t),$$

où $p(t)$ est le prix de l'output produit, F la fonction de production, K le stock de capital, $w(t)$ le salaire et L la main d'oeuvre.³ Nous dénotons par $R(t)$ le revenu net au temps t qui est donné par

$$R(t) = (1 - u(t))\pi(t) + u(t) \int_0^\infty D(x, t-x)p_I(t-x)I(t-x) dx - (1 - k(t))p_I(t)I(t),$$

où on dénote par $\pi(t)$ le profit avant taxation au temps t , $u(t)$ la taxe corporative, $D(x, t-x)$ le taux de dépréciation par dollar d'investissement taxé âgé de x années au temps $t-x$, $p_I(t)$ le prix de l'investissement, $I(t)$ l'investissement et $k(t)$ le taux de taxation sur l'investissement. Finalement, une firme doit maximiser la valeur présente de son revenu futur net après taxe $V(0)$ donnée par

$$V(0) = \int_0^\infty R(t)e^{\int_0^t r(s)ds} dt.$$

Les marchés sont parfaitement compétitifs impliquant ainsi que les firmes considèrent les prix comme exogènes. Jusqu'à présent, rien de nouveau n'est apparu, cependant l'idée de coût de remplacement se matérialise dans la contrainte d'optimisation suivante.

$$\frac{\partial K}{\partial t} = \psi(I, K, t) - \delta K,$$

où δ est le taux de dépréciation physique du capital et ψ une fonction exprimant le pourcentage de l'investissement qui se transforme en capital. Deux hypothèses importantes sont faites sur ψ soit que cette dernière est non décroissante et concave. L'introduction de la fonction ψ est dû à Uzawa [69] et permet de tenir compte du coût d'installation des unités d'investissement, le coût d'ajustement.⁴ La figure 2.1 représente la fonction ψ .

3. Hayashi dans [38] utilise $N(t)$ le vecteur des inputs nécessaire au lieu de $L(t)$. Nous utilisons une version simplifier qui n'altère en rien les conclusions obtenues.

4. Lucas [57], Gould [36] et Treadway [68] proposent une approche différente faisant intervenir la fonction

On constate que la courbe décroît très rapidement lorsque l'investissement devient négatif.

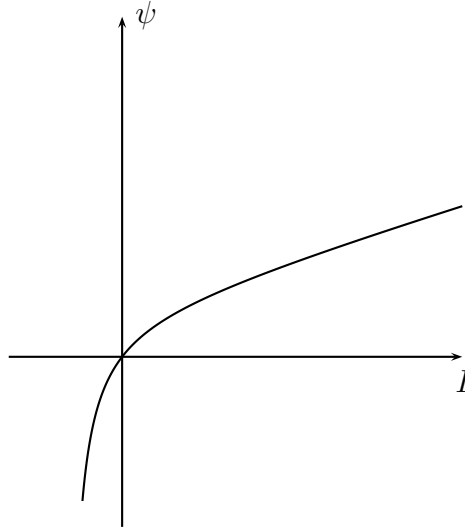


FIGURE 2.1 – Fonction ψ

Cette caractéristique s'explique par le fait que ψ reflète l'irréversibilité de l'investissement. Ainsi, le problème de maximisation de la firme est donné par

$$\begin{aligned} \max \quad & \int_0^{\infty} R(t) e^{-\int_0^t r(s) ds} dt \\ \text{tel que} \quad & \frac{\partial K}{\partial t} = \psi(I, K, t) - \delta K. \end{aligned}$$

Une réécriture de $V(0)$ par deux changements de variables successifs permet d'obtenir

$$V(0) = \int_0^{\infty} [(1 - u(t))\pi(t) - (1 - k(t) - z(t))p_I(t)I(t)] e^{-\int_0^t r(s) ds} dt \quad (2.1)$$

$$+ \int_0^{\infty} u(t) \left[\int_{-\infty}^0 D(t - \nu, \nu) o_I(\nu) I(\nu) d\nu \right] e^{-\int_0^t r(s) ds} dt \quad (2.2)$$

de profit de la manière suivante

$$\pi(t) = p[F(K, L, t) - G(I, K, t)] - wL,$$

où G dénote la fonction de coût d'installation de l'investissement qui dépend du ratio entre I et K . Il est important de noter que dans ce cas la fonction G est telle que $G_I > 0$ et $G_{II} > 0$.

où

$$z(t) = \int_0^\infty u(t+x)D(x,t)e^{-\int_0^x r(t+s)ds}dx.$$

Le deuxième terme de l'équation 2.1 représente la valeur actuelle et future des déductions de taxe attribuable à l'investissement passé. Suivant la notation de Hayashi, nous notons ce terme $A(0)$. Avec cette nouvelle expression, les variables de choix du problème d'optimisation sont I et L .

Afin de résoudre ce problème d'optimisation en temps continu, nous appliquons le principe du maximum de Pontryagin [64]. Nous remarquons que le terme $A(0)$ n'intervient en aucun cas dans le processus d'optimisation. Avec cette remarque, l'Hamiltonien est donné par

$$H(I, L, t, \lambda) = [(1 - u(t))\pi(t) - (1 - k(t) - z(t))p_I(t)I(t)]e^{-\int_0^t r(s)ds} + \lambda[\psi(I, K, t) - \delta K].$$

Les conditions du premier ordre sont données par

$$\frac{\partial H}{\partial L} = (1 - u(t))\frac{\partial \pi(t)}{\partial L} = 0 \quad (2.3)$$

$$\frac{\partial H}{\partial I} = [(1 - u(t))\pi(t) - (1 - k(t) - z(t))p_I(t)]e^{-\int_0^t r(s)ds} + \lambda\frac{\partial \psi(I, K, t)}{\partial I} = 0 \quad (2.4)$$

Notons que l'équation (2.4) est différent de l'équation 8 de [38].⁵ En dérivant l'équation 2.4 par rapport au temps et utilisant l'identité fournie par cette dernière, nous obtenons

$$\frac{\partial \lambda}{\partial t} = (r + \delta + \psi_K)\lambda - (1 - u)\pi_K. \quad (2.5)$$

La condition de transversalité est donnée par

$$\lim_{t \rightarrow \infty} \lambda(t)K(t)e^{-\int_0^\infty r(s)ds} = 0$$

Comme à l'habitude, l'interprétation de l'équation (2.3) est une condition de productivité

5. Le terme $e^{-\int_0^t r(s)ds}$ est manquant mais n'affecte en rien les calculs.

marginale. L'équation (2.4) stipule que le coût marginal d'installation d'une nouvelle unité d'investissement est égal au bénéfice de cette unité. Il est plus facile de voir cette interprétation en réécrivant l'équation (2.4) sous la forme suivante

$$(1 - k)P_i + (1 - \psi_I)\lambda = \lambda + zp_I$$

où zp_i s'interprète comme la valeur de la déduction de taxe obtenue pour une unité d'investissement.

On définit le q -marginal de Tobin par

$$q = \frac{\lambda}{p_I}$$

et le q -moyen de Tobin par

$$h = \frac{V(0)}{p_I K}.$$

Les équations de premier ordre peuvent donc être écrites comme suit.⁶

$$\frac{q}{1 - k - z} = \frac{1}{\psi_I} \tag{2.6}$$

$$\dot{q} = \left(r + \delta - \frac{\dot{p}_I}{p_I} - \psi_K \right) q - (1 - u) \frac{\pi_K}{p_I} \tag{2.7}$$

L'équation 2.6 permet de déterminer l'investissement optimal en fonction de K et $\frac{q}{1-k-z}$.⁷ Fait important, si les firmes connaissent q alors la seule connaissance de la fonction ψ , qui est la fonction de coût d'installation du capital, permet de déterminer directement le stock de capital optimal.

6. Notons que nous utilisons ici les équations de [38] qui ne tiennent pas compte du facteur $e^{-\int_0^t r(s)ds}$.

7. Habituellement, le quotient $\frac{q}{1-k-z}$ est noté par \tilde{q} et nommé q -modifié dans la littérature.

2.1.2 Extension : L'incertitude et l'irréversibilité

La manière dont nous venons de définir q est l'un des avantages de cette théorie, i.e. de solides bases théoriques en microéconomie. Il est ainsi facile en considérant le comportement d'optimisation des firmes d'y introduire l'incertitude. En réalité ce n'est qu'en 1994 dans [1] qu'une approche unificatrice est présentée afin d'unir à la fois l'incertitude et un coût d'investissement fixe. Le coût d'investissement fixe en interaction avec l'incertitude met en relief un facteur important tenu à l'écart de la théorie présentée à la section précédente qui réside dans la perception des entreprises de l'irréversibilité. Dans le modèle déterministe précédent, l'irréversibilité est sûre et connue de la firme qui sais dors et déjà quelle partie de son investissement sera engloutie en coût d'ajustement. Or le simple fait de ne plus savoir exactement ce ratio peut remettre en question tout le processus décisionnel d'optimisation de la firme. Une comparaison simple est de considérer l'investissement de la firme comme l'analogie de l'épargne pour le consommateur. Ainsi, face à une incertitude, les appréhensions du taux d'intérêt affecteront son taux d'épargne actuel tout comme la firme face à une augmentation du pourcentage de l'investissement nécessaire en coût d'ajustement sera plus enclin à diminuer cet investissement et d'attendre possiblement pour maximiser son utilisation.

Cette particularité est développée de façon théorique par Abel et coll. dans [1] où trois scénarios d'investissement sont possibles pour les firmes :

1. investissement négatif,
2. investissement nul,
3. investissement positif.

Le troisième scénario distingue cet article des autres parues avant celui-ci. Effectivement, les articles traitant des coûts d'ajustement dans le cadre de la théorie de l'investissement envisagent seulement le premier scénario, alors que ceux portant sur l'irréversibilité de l'investissement ne considèrent que les deux premiers scénarios. Certaines hypothèses sont faites sur l'investissement à savoir que celui-ci est une fonction croissante de q . De plus, il est important de noter que le coût d'ajustement est une fonction convexe et donc introduite dans la fonction de profit au lieu de suivre l'approche utilisée à la section précédente. Pour ainsi

discerner les trois possibilités d'investissement d'une firme, Abel et coll. relie ces comportements avec des valeurs critiques de q . Plus précisément,

$$I(q, K) = \begin{cases} < 0 & \text{for } q < c_I(0, K)^- \\ = 0 & \text{for } c_I(0, K)^- \leq q \leq c_I(0, K)^+ \\ > 0 & \text{for } q > c_I(0, K)^+ \end{cases}$$

où $c_I(I, K)$ est la fonction de coût d'ajustement augmenté et où $c_I(0, K)^-$ et $c_I(0, K)^+$ dénote la dérivé partielle à gauche et à droite au point $I = 0$.⁸ Avec cette notation, $c(0, K)$ est vu comme le coût fixe d'investissement et est supposé strictement positif. Avec cette approche les auteurs de [1] montrent que le choix de ne pas investir n'est pas un cas dégénéré.

De plus, Abel et coll. essaient de palier à une des plus grandes lacunes de la Théorie- q , i.e. le fait que q n'est pas observable. Pour ce faire, ils expriment q comme la valeur de la firme divisée par le stock de capital dans un premier temps et dans un second exemple comme une fonction de l'output, de la fonction de production ainsi que du taux d'intérêt. Il est à noter que pour ce faire la fonction de production est supposée homogène de degré 1 et constitue un point crucial dans leur démonstration. Les marchés sont également en parfaite compétition.

2.1.3 Problème dynamique des données

Comme nous l'avons mentionné précédemment, l'un des gros désavantages de la Théorie- q est que l'acteur clé, q , n'est pas observable. Ainsi, les tests empiriques doivent utiliser une proxy qui est habituellement le q -moyen. Avec cette mesure, le modèle économétrique a de solides incapacités à tenir compte des variations temporelles de l'investissement au niveau agrégé [66]. Aucune raison théorique acceptable ne justifie l'ajout de variables, telle la contrainte de liquidité, qui sont significatives dans plusieurs études. Cette incapacité d'expliquer convenablement les données est un autre aspect négatif de cette théorie. Blunbel et coll. [17] argumente que l'utilisation d'une variable proxy pour q fausse les résultats étant

8. L'expression donnée pour $I(q, K)$ suppose que la firme choisie d'investir. Cette hypothèse se traduit par $\nu = 1$ dans [1]. Il est certain qu'il y a une relation de dépendance avec le choix d'investir et $I(q, K)$ dans le processus d'optimisation.

donnée la dynamique de la variable *proxi*, ou du moins de leurs mesures.

Blunbel et coll. traite la structure dynamique du problème de façon minutieuse en utilisant un estimateur obtenu par la méthode des moments généralisée. Les auteurs notent que le choix de cet estimateur se veut un outil de correction pour la possible présence de l'hétéroscédasticité. Cette étude utilise des microdonnées de 532 firmes se situant au Royaume-Uni de 1975 à 1986. Blunbel et coll. mentionne que l'utilisation de microdonnées est préférable afin de tester le modèle économétrique Tobin- q . Cette affirmation se base sur le fait que q est propre à chacune des firmes, à la limite à chacune des industries. Nous abondons également dans ce sens en ajoutant que le q -moyen utilisé avec les données agrégées ne tient pas compte des particularités propres de chacune des industries pouvant faire face à des coûts d'ajustement fort différents pouvant dépendre directement de la technologie utilisée. Cet argument est confirmé avec le deuxième cas présenté dans [1] où le q est obtenu comme une fonction de plusieurs variables, dont la fonction de production, qui internalise la technologie propre à chacune des industries.

Comme prévu, Blunbel et coll. confirme par leurs résultats que les paramètres estimés sont grandement sensibles aux paramètres du modèle, étant reliés directement à la fonction de production utilisée, mais également à la spécification dynamique de la mesure et aux erreurs de mesure du q -moyen. Ils concluent que la variable q -moyen est significative, mais avec un faible coefficient proposant ainsi que le taux d'investissement à court terme ne répond que très faiblement aux variations des valeurs du capital. Une autre conclusion obtenue est que les restrictions imposées sur leur modèle afin de calculer le q -moyen ne sont pas satisfaisantes, car ce dernier ne reflète pas convenablement de façon empirique tous les déterminants importants de la prise de décision relative à l'investissement comme le propose la théorie. Nous mentionnons ici qu'une telle conclusion est plus qu'évidente à la lumière des Propositions 1 et 2 de [1] démontrant l'équivalence entre le q -marginal comportant toute l'information nécessaire à la firme et le q -moyen. Finalement, nous mentionnons que puisque le stock de capital est nécessaire à l'approche utilisée par Blunbel et coll., il est plus que probable que des erreurs de mesure soient présentes dans les données de stock de capital.

Cette conclusion se base sur [30]. Nous terminons en mentionnant que Blunbel et coll. estiment avec les données qu'ils ont à leurs dispositions les taxes pour chacune des entreprises afin de calculer le q -moyen. Il va s'en dire que cette méthodologie permet de tenir compte du niveau de taxation des entreprises, mais empêche d'isoler son effet direct sur l'investissement. Ce point est une autre critique du modèle Tobin q . À notre avis, ce désavantage en est un insurmontable, du moins du côté empirique lorsqu'il est question d'étudier le comportement de l'investissement dans son ensemble tout en souhaitant déterminer quelles sont les facteurs les plus importants pour l'investissement. Seul un q observable, pourrait réellement remédier à cette lacune permettant de mieux saisir l'impact de chacun des facteurs sur q qui par la suite influence directement le choix des firmes toujours dans le cadre du modèle de Tobin- q .

2.2 Théorie de la gravité

Cette section porte sur le modèle de gravité qui est largement utilisée dans le cadre de l'étude empirique de l'investissement direct étranger. Premièrement, nous présentons les grandes lignes de cette théorie ainsi que les motivations qui poussèrent à l'utiliser pour l'étude de l'investissement direct étranger. Ensuite, nous illustrons les avantages et les inconvénients que comporte le modèle empirique lié à cette théorie en puisant dans plusieurs articles empiriques.

2.2.1 L'intuition derrière la théorie de la gravité et ses grandes lignes

Développé dans le cadre de la théorie des échanges commerciaux entre les pays, le modèle de gravité est reconnu pour sa grande performance avec les données. Une façon simple et intuitive d'amorcer son étude est de faire une analogie avec la magnifique et très célèbre Loi de la Gravitation de Newton. Voici donc ce que l'on pourrait appeler une intégration économique de cette fameuse loi.

Le flux d'échanges commerciaux entre deux pays s'intensifie à mesure que la distance économique les séparant diminue, et ce en tenant compte de leur poids économique respectif.

Il est certain que cette phrase est vague, mais résume malgré tout très bien tous les points essentiels de cette théorie. Pour être plus précis, on peut définir la distance économique par la distance géographique entre deux pays, utilisé fréquemment dans les études empiriques, et leurs poids par le PIB respectif des pays considérés. Ainsi, mathématiquement on obtient

$$PX_{AB} = \beta_0 (PIB_A)^{\beta_A} (PIB_B)^{\beta_B} (DIST)^{\beta_3} \quad (2.8)$$

où PX_{AB} est la valeur des échanges commerciaux entre le pays A et B et $DIST$ la distance séparant ces deux pays. Un fait économique notable de cette simple équation est qu'elle donne explicitement comment la distance exerce une friction sur les échanges commerciaux bilatéraux. De plus, l'équation (2.8) ne tient compte que des échanges effectués entre A et B . Nous traitons ce point plus en détail plus tard.

Bien que l'équation (2.8) puisse être sensée et empiriquement très attrayante, il n'en reste pas moins que seule une fondation microéconomique solide peut justifier son utilisation. Ce point était une lacune importante du modèle de gravité et ce n'est que récemment qu'elle a été partiellement résolue dans le cadre des échanges commerciaux entre pays. Une revue de la théorie de la gravité est présentée par Anderson dans [2] où il établit des bases microéconomiques solides afin d'introduire le modèle de gravité de manière très rigoureuse. Cette approche utilise l'équilibre général où une simplification, nommé modularité⁹, sur la structure de l'équilibre lui-même. Cette propriété se décrit comme "subordinating the economic determination of equilibrium distribution of goods within a class under the superstructure determination of the distribution of production and expenditure between classes of goods." [2] L'avantage d'une telle simplification réside dans le fait que l'on peut considérer seulement l'inférence des coûts de distorsion causé par la distribution sans avoir à expliquer comment la fourniture de la demande totale est déterminée pour chacun des emplacements.

Plusieurs recherches empiriques mentionnent le fait que la distance entre 2 pays est significative pour expliquer l'investissement direct étranger entre 2 pays. Pour cette raison,

9. Anderson et Van Wincoop dans [3] utilise le terme propriété de séparabilité du commerce au lieu de modularité.

le modèle de gravité est appliqué au contexte de l'investissement direct étranger (FDI) par analogie sans nécessairement avoir de véritable fondement microéconomique. Cette critique tend à s'estomper depuis les travaux de Helpmann et coll. [40] dans lequel une modélisation de l'interaction entre les ventes de firmes étrangères, ici vue comme une mesure du FDI, et les exportations dites horizontales est faite dans le contexte d'un modèle d'équilibre général. Ce modèle est amélioré par Kleinert et Toubal [45] en introduisant les coûts fixes de l'investissement qui croient avec la distance entre les pays. Kleinert et Toubal obtiennent également une équation de gravité pour l'investissement dit vertical.¹⁰

2.2.2 La présence de 0

Malgré la mondialisation et l'ouverture des marchés, il est évident qu'un pays n'investit pas dans tous les autres pays de la planète. Cette simple évidence a cependant de lourdes conséquences sur le modèle de gravité. En fait, le modèle lui-même en semble être la cause. Une simple réécriture de l'équation 2.8 permet de voir tout de suite pourquoi.

$$\ln(FDI) = \ln(\beta_0) + \beta_A \ln(PIB_A) + \beta_B \ln(PIB_B) + \beta_3 \ln(DIST_{AB}) \quad (2.9)$$

Les flux d'investissements nuls *observés* peuvent être interprétés de plusieurs manières.

1. Il n'y a effectivement pas d'investissement entre les pays concernés.
2. L'investissement est trop faible et donc négligé par l'observateur ou encore des erreurs de mesures introduisent un biais vers le bas avec comme résultat un investissement nul.

Une première solution est de ne pas tenir compte des observations qui présentent un investissement nul. Il est certain que si le 0 observé est réellement un 0, alors cette observation n'apporte rien à l'analyse (dans le cadre du modèle de gravité). Cette approche est utilisée depuis fort longtemps pour traiter ce problème. Ledyeva et Linden [47] en font l'utilisation pour étudier les facteurs déterminants l'investissement direct étranger entrant dans les régions de la Russie. Pour ce faire, ils utilisent des données portant sur l'investissement des 5 plus importants pays investisseurs en Russie dans plus de 76 régions en 2002. Cette analyse transversale utilise deux méthodes d'estimation soit la méthode des moindres carrés

10. Voir [2, section 6.3] pour un résumé plus complet des récentes découvertes.

(OLS) où effectivement toutes les observations affichant un investissement nul sont écartées. Par la suite le modèle d'estimation logit et probit sont utilisés. Il faut noter ici que l'investissement direct étranger consiste à compter le nombre de firmes étrangères par région. Ledyaeva et Linden motivent ce choix directement avec leur question de recherche portant sur les facteurs déterminants de l'investissement. Ainsi, l'utilisation d'une mesure donnant directement la valeur totale de l'investissement contient possiblement l'effet d'attraction de ressources naturelles et ainsi interfère avec l'estimation de l'impact d'autres facteurs importants. Les auteurs affirment également que le nombre de firmes capture bien l'intensité de l'investissement direct étranger. Le fait d'affirmer que le nombre de firmes capture bien l'intensité de l'investissement direct étranger suppose que ces firmes sont toutes homogènes ne laissant pas la possibilité qu'une firme seule dans une région puisse investir à elle seule davantage que toutes celles réunies dans une autre région. Ce genre de situation peut se produire lorsque l'on compare les investissements de compagnies minières avec des firmes dans le secteur manufacturier des vêtements. Il est donc de l'avis de l'auteur de la présente étude que cette mesure est inadéquate. Un fait marquant de cette étude est que les estimations par OLS, par logit et par probit produisent toutes les mêmes résultats qualitativement. La comparaison quantitative de ces régressions est difficile de par la nature des méthodes d'estimation et aussi par le fait que Ledyaeva et Linden n'incorporent pas les mêmes variables binaires dans leurs différentes régressions.

Le problème d'erreur de mesure quant à lui est plus problématique. Une solution proposée est encore une fois d'ignorer ces données. La justification de cette approche est que si jamais une erreur de mesure provoque un investissement nul alors, l'investissement réel est très faible et donc de peu d'importance pour l'analyse. Ce raisonnement contient deux failles. La première sous-entend que l'erreur de mesure est elle-même petite. Or il est très difficile de pouvoir juger de l'importance d'une erreur de mesure, car c'est une erreur de mesure inobservable! La deuxième est qu'il est possible que cette erreur de mesure bien que possiblement petite soit causée par l'observateur lui-même proposant donc un biais de sélection. Cette éventualité est étudiée dans [37]. Pour ce faire, Cranfield et coll. utilisent la méthode de sélection de Heckman avec des données portant sur les échanges commerciaux

entre 52 pays allant de 1990 à 2000. Leurs estimations considèrent 46 produits provenant du secteur de l'agriculture supposés tous différents aux yeux du producteur et du consommateur. Ils suivent étroitement la méthode de Berstand [10] [11] et supposent que les préférences sont de type Linder. Ils introduisent également la possibilité de coût de production fixe. Cette hypothèse est cruciale, car elle permet de justifier théoriquement la présence d'échanges commerciaux nuls entre deux pays. Leurs régressions comportent près de 130 paramètres à estimer en incluant les effets fixes de pays et de périodes. On constate qualitativement, grâce à leurs résultats, que la méthode de sélection de Heckmann des moindres carrés ne diffère que très peu. En effet, seuls trois paramètres ont changé de signe illustrant un changement d'effet important. La véritable importance de la sélection de biais se situe dans l'intensité de l'impact des facteurs qui déterminent les échanges commerciaux. La variation observée peut aller jusqu'à 20%. Cranfield et coll. interprète par la suite les estimés en terme d'élasticité démontrant ainsi comment une variation de 20% affecte grandement les échanges commerciaux.

La même problématique est abordée par Wagle dans [70] où l'investissement direct étranger est considéré plutôt que les échanges commerciaux. Wagle utilise également la méthode de sélection de Heckman afin de tenir compte du possible biais de sélection. Il est intéressant de noter qu'il porte une attention particulière à une autre source de biais, l'hétérogénéité des firmes. Ce biais s'amplifie lorsque les firmes n'ont pas la capacité de défrayer les coûts fixes de l'investissement. Ce point le distingue de [37] où aucune mention de l'hétérogénéité n'est faite. Encore ici, le fait d'avoir un coût fixe à l'investissement rend possible l'observation de 0. La présence de l'hétérogénéité se justifie par la différenciation à l'intérieur des industries en terme de productivité d'utilisation d'input ou encore de salaires versés.¹¹ Basé sur le modèle de gravité découlant de Anderson et Van Wincoop [3], Wagle examine comment les politiques spécifiques aux pays relatifs à l'investissement direct étranger ainsi que les institutions affectent l'attraction de FDI. Il est à noter ici que les régressions sont faites en cross section avec des pays membres et non-membres de l'OCDE entre 2007 et 2008. Ces résultats démontrent bien l'existence d'erreur de sélection. Fait étonnant, la qualité des institutions, considérée dans les régressions, est non significative. Encore plus étonnant, le niveau de sco-

11. Melits [59] expose des raisons pour lesquels seul une partie des firmes investissent à l'étranger.

larité est également non-significatif. Ces deux résultats semblent suggérer que le contexte social ainsi que la qualité de la main d'oeuvre ne sont pas considérés par les investisseurs lorsque l'investissement est fait dans des économies assez développées.

2.2.3 Le modèle de gravité et la taxation corporative

Un des avantages du modèle de gravité par rapport à celui de Tobin- q est la possibilité de tester directement quels sont les impacts de la taxation corporative sur l'investissement direct étranger. En effet, la Théorie- q incorpore dans le calcul du q les effets des taxes ainsi que d'autres facteurs déterminants de l'investissement. Il faut voir que dans le cadre d'un modèle de gravité, l'investissement allant du pays A vers le pays B est la variable dépendante et que les variables explicatives sont celles du pays hôtes de l'investissement ainsi que quelques variables du pays d'où provient l'investissement rendant ainsi possible l'ajout d'une variable capturant l'effet de la taxation des entreprises.

Un article étudiant en détail cet aspect est celui de Benassy-Queré et coll. [9]. Benassy-Queré et coll. utilisent différentes mesures de taxation telles que le taux de taxation nominal, le taux moyen effectif, le taux marginal effectif ainsi que le taux apparent effectif sur des données de panel de 11 pays membres de l'OCDE entre 1984 et 2000. Leur premier résultat montre que le niveau de taxation est significatif. Cependant, l'utilisation des différentes mesures de taxation n'a de l'influence que sur la valeur du coefficient estimé, c'est à dire des semi-élasticité, plutôt que sur la significativité de ce dernier comme le démontre le Tableau 2.1 extrait de [9]. Tous les valeurs estimées présentées dans le Tableau 2.1 sont significatives

Mesure du taux de taxation	Valeur du coefficient estimé
Nominal	-4,23
Moyen effectif	-2,89
Marginal effectif	-9,4
Apparent effectif	-4,22

Tableau 2.1 – Résultats de l'estimation de l'équation de base de [9]

à un niveau de 99%. Ainsi, on y voit que toutes les taxes ont un impact négatif sur l'investissement de manière à ce que l'augmentation d'un point de pourcentage du taux nominal,

moyen effectif, marginal effectif et apparent effectif de taxation provoque une diminution de 4,23 %, 2,89 %, 9,4% et 4,22% de l'investissement respectivement

Une caractéristique qui différencie cette étude des autres qui étudient les impacts de la taxation sur l'investissement avec le modèle de gravité est résumée dans le passage suivant : "Indeed, because investors arbitrate not only among foreign locations but also between each foreign location and domestic investment, tax differentials rather than tax levels should be their variable of decision." [9] Pour tenir compte de cette réalité, les auteures étudient également la différenciation du niveau de taxation entre les pays. En séparant les différenciations de taxe par leurs signes, i.e. soit positif ou négatif dépendant de la paire ordonnée de pays investisseur et hôte, Benassy-Queré et al trouve que le coefficient d'une différenciation de taxe positive est plus élevé en valeur absolue que celui estimé pour les différenciations de taxe négatives. Ce résultat signifie que l'investissement direct étranger entrant est plus affecté par un taux de taxation élevé. Ceci implique aussi que l'entrée de capital d'investissement est moins sensible aux incitations fiscales à investir comparativement aux mesures pénalisant l'investissement. Ce résultat est d'une grande importance dans l'élaboration de politiques fiscales concernant l'investissement étranger, car il suggère que les coupures de taxe seront beaucoup plus efficaces dans des pays qui ont un haut taux de taxation comparativement à ceux avec un faible taux qui effectue la même coupure.

Finalement, un autre fait saillant de [9] est que les auteurs testent la possibilité d'un impact non linéaire des niveaux de taxation, plus exactement sur la différence des niveaux de taxation entre les pays. Pour ce faire, ils introduisent le cube de la variable du niveau de taxe dans les régressions. Dans cette analyse, les auteurs tiennent également compte de l'accès au crédit par les firmes. Lorsque le crédit est pris en considération, la différenciation du niveau de taxation a un impact non linéaire sur l'investissement alors que l'impact n'est que linéaire dans le cas contraire. Cette asymétrie suggère une autre façon de stimuler l'investissement direct étranger pour les gouvernements. Ce point est non-négligeable dans un contexte où une "course vers le bas" du niveau de taxation des entreprises est déjà commencée!

2.3 Économétrie spatiale

Dans cette section, nous présentons en un premier temps les deux problèmes fondamentaux de l'économétrie spatiale. Pour ce faire nous nous basons sur le livre de LeSage [53]. Nous abordons ensuite les différentes spécifications du modèle spatial en illustrant à l'aide d'articles empiriques les avantages et inconvénients de chacun d'entre elles. Nous traitons aussi des problèmes d'estimation liés à chacune de ses spécifications.

2.3.1 Problème généraux de l'économétrie spatial

Les recherches faites en sciences régionales utilisent des données recueillies auxquelles une localisation dans l'espace est ajoutée. Ainsi, l'économétrie spatiale traite de la manière dont doit être inclus cet aspect supplémentaire dans les données et la spécification du modèle afin que celui-ci soit pris en considération lors des régressions et de l'analyse subséquente. La théorie de la gravité est un exemple de théorie qui prend en considération la composante spatiale lorsque la distance utilisée est géographique. Au début des années 2000, plusieurs chercheurs mentionnent que le modèle de gravité ne tient pas compte dans le processus décisionnel des autres emplacements possibles de l'investissement puisque seule la relation bilatérale entre deux pays est considérée. Dans la littérature, on nomme cette interaction *l'effet du troisième pays* [6] [35]. L'apparition de cette nouvelle terminologie est due à l'article pionnier de Coughlin et Segev [29] portant sur l'investissement des États-Unis dans les différentes provinces et régions de la Chine incorporent explicitement comme variable explicative une variable déterminant si certaine région sont voisine les une des autres.¹² Dans leur article, Coughlin et Segev mentionnent que : " Our analysis uses the location of observations to test for the existence of two potential econometric problems - spatial heterogeneity and spatial dependence - that raise doubts about the results from ordinary least squares regressions." [29]. Ces deux problèmes sont fondamentaux en économétrie spatiale, car ceux-ci violent les hypothèses du théorème classique de Gauss-Markov. Ainsi, les estimés obtenus par la méthode des moindres carrés sont biaisés.

12. Deux régions peuvent être définies comme voisine s'il partage un sommet ou une partie de frontière commune. Cet aspect est traité plus en détail dans [53, section 1.4]

L'hétérogénéité spatiale fait référence à une variation reliée à la localisation de l'observation. Nous formalisons ceci de la manière suivante.

$$y_i = \beta_i X_i + \epsilon_i, \quad i = 1, \dots, n, \quad (2.10)$$

où β_i le vecteur des paramètres à estimer, X_i les vecteurs des variables explicatives et i la localisation de l'observation. L'indice i associé à β signifie que pour chaque localisation, la relation étudiée est différente. Il est normal de penser que la relation est dépendante de sa localisation. Par exemple, il est peu probable que la main d'oeuvre de la Corée du Nord réponde de la même manière que celle des États-Unis face à l'ajout d'un impôt forfaitaire. Cependant, l'estimation de l'équation (2.10) est impossible dû à un problème de degré de liberté. La solution réside dans une spécification du modèle qui tient compte de cette variation spatiale. Il est certain qu'une étude approfondie des données et de la question d'intérêt mènent à différentes spécifications.

Formellement, la dépendance spatiale s'énonce comme la dépendance des observations face à leurs localisations. On écrit

$$y_i = f(y_i) \quad i = 1, \dots, n$$

où y_i est l'observation à la localisation i . Cette relation de dépendance entre les observations via la dimension spatiale viole l'une des hypothèses classiques de Gauss-Markov stipulant que les variables explicatives sont fixées dans un échantillon répété. Il est vrai que cette hypothèse peut être évitée, cependant la dépendance spatiale cache en réalité une faille plus profonde. Pour illustrer notre propos, nous utilisons un exemple de LeSage [53, section 1.2] qui est le suivant. Le mobilité de la main-d'oeuvre se veut être la possibilité d'un travailleur de trouver de l'emploi dans une région différente de celle où il réside. Cependant, si le taux de chômage est fonction du lieu de résidence, alors ce dernier ne tient pas compte de cette mobilité. Par conséquent, en associant une dimension spatiale au taux de chômage mesuré par le lieu de résidence, nous introduisons une erreur de mesure. C'est cette erreur de mesure

qui, contenue dans la variable explicative, provoque la dépendance spatiale. En effet, les erreurs de mesure sont toutes entre elles dépendantes selon leurs régions voisines. De façon générale, il est supposé que l'importance de cette erreur de mesure diminue avec la distance séparant les observations.

L'outil de prédilection pour traiter la dépendance spatiale est une matrice W qui permet de relier entre elles les régions voisines selon leurs poids respectifs. Aucune définition définitive n'est possible pour W étant donné que la matrice dépend de la définition du terme "voisin" de même que de la définition du poids qui peut grandement varier pour tenir compte du problème. La seule caractéristique commune est que celle-ci possède des rangées normalisées, i.e. la somme des composantes de chaque rangée est égale à 1. Nous présentons ici une généralisation de la définition utilisée par Blongien et coll. dans [15]. Soit W une matrice carrée diagonal par bloc de dimension n . Soit $t \in [1, T]$ l'indice de temps allant de 1 à T tenant ainsi compte de la structure possible des données de panel. Nous définissons W de telle sorte que chaque bloc de la diagonale capture la dépendance spatiale d'une seule période de temps. On définit donc un bloc diagonal W_t par

$$W_t = \begin{bmatrix} 0 & w_t(d_{12}) & w_t(d_{13}) & \cdots & w_t(d_{1k}) \\ w_t(d_{21}) & 0 & w_t(d_{23}) & \cdots & w_t(d_{2k}) \\ \vdots & \ddots & & \ddots & \\ w_t(d_{n1}) & \cdots & \cdots & \cdots & 0 \end{bmatrix},$$

où $w_t(d_{i,j})$ est une fonction de poids décroissante de la distance $d_{i,j}$ entre les régions i et j .

Par exemple,

$$w_t(d_{i,j}) = \begin{cases} \frac{\min_{(l,k) \neq (i,j)} d_{lk}}{d_{ij}} & i \neq j \\ 0 & i = j \end{cases}$$

Avec cette définition, un poids égal à 1 est attribué à la plus petite distance séparant deux pays et sert ainsi de références pour de futures comparaisons.¹³ De plus, les 0 sur la diagonale assurent qu'aucune variable dépendante ne peut être expliquée par elle-même. Il suffit par la suite de normaliser les rangées de la matrice W .

13. Cette définition peut être facilement modifiée afin de choisir la distance de référence.

2.3.2 Modèle SAR

Dans cette sous section, nous présentons premièrement le modèle autorégressif d'ordre 1 afin d'illustrer les problématiques de l'estimation des moindres carrés. Par la suite, nous introduisons le modèle autorégressif spatial standard (SAR). Nous relierons les différentes motivations de l'investissement des firmes avec les signes envisagés des paramètres à estimer du modèle SAR. Nous examinons certains articles de la littérature empiriques dans le but de connaître les forces et les faiblesses du modèle. Finalement, nous nous attardons à un problème économétrique qui consiste à estimer le modèle SAR avec des données de panel non balancé.

Estimation biaisée des moindres carrés

Tel que mentionné précédemment, la méthode des moindres carrés fournit des estimés biaisés lorsqu'il y a de la dépendance spatiale. Pour démontrer ce fait, nous utilisons un modèle simple autorégressif spatial d'ordre 1 donné par

$$y = \rho W y + \epsilon, \quad \epsilon \sim N(O, \sigma^2 I_n) \quad (2.11)$$

où W est la matrice des poids spatiaux. Ainsi l'estimateur des moindres carrés est donné par

$$\hat{\rho} = (y'W'Wy)^{-1} y'W'y.$$

Pour vérifier que cet estimateur est biaisé, il suffit de calculer son espérance.

$$\begin{aligned} E[\hat{\rho}] &= E \left[(y'W'Wy)^{-1} y'W'y \right] \\ &= E \left[(y'W'Wy)^{-1} y'W'(\rho W y + \epsilon) \right] \\ &= E \left[(y'W'Wy)^{-1} y'W'\rho W y \right] + E \left[(y'W'Wy)^{-1} y'W'\epsilon \right] \\ &= \rho + E \left[(y'W'Wy)^{-1} y'W'\epsilon \right]. \end{aligned}$$

L'astuce habituelle des moindres carrés est de conditionner sur la variable dépendante, nommons-la X pour l'instant, et d'utiliser la non corrélation entre X et ϵ .¹⁴ Or ici, même si l'on conditionnait sur la variable y celle-ci est corrélée avec ϵ d'où

$$E \left[(y'W'Wy)^{-1} y'W'\epsilon \right] \neq 0$$

et donc

$$E[\hat{\rho}] \neq \rho.$$

De plus, une réécriture de l'équation (2.11) obtenue comme suit

$$\begin{aligned} y &= \rho W y + \epsilon \\ y - \rho W y &= \epsilon \\ (I - \rho W)y &= \epsilon \\ y &= (I - \rho W)^{-1} \epsilon \end{aligned}$$

permet de déduire que l'estimateur des moindres carrés n'est pas consistant.¹⁵ En effet, Anselin [4] obtient que

$$plim \frac{1}{N} y'W'\epsilon = plim \frac{1}{N} \epsilon'(I - \rho W)^{-1} \epsilon \neq 0$$

montrant ainsi l'inconsistance de l'estimateur.

14. Noter que cette explication diffère de celle donnée dans [53] où l'on traite un cas particulier, i.e. le cas où X est prédéterminé dès le départ.

15. Ici, certaines hypothèses sur la matrice W doivent être faites ou sur le paramètre ρ afin de pouvoir assurer que $(I - \rho W)^{-1}$ existe.

SAR

Le modèle spatial autorégressif souvent appelé "spatial lag model" dans la littérature est donné par

$$y = \rho W y + X\beta + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I_n), \quad (2.12)$$

où X est un vecteur de variables explicatives contenant une variable capturant l'importance du marché du pays hôte.

Un avantage marqué du modèle SAR est que les effets marginaux produits par les variations de variables spécifiques à un pays sont facilement calculables. En effet, les effets marginaux affectent la variable dépendante, mais également les pays voisins. Dans ce cas on parle de retombées (Spillover). LeSage et Pace dans [55] proposent un manière de calculer les effets directs et indirects.¹⁶ Voici en quoi elle consiste. Partant de l'équation (2.12), on obtient

$$\begin{aligned} (I - \rho W)y &= X\beta + \epsilon \\ y &= (I - \rho W)^{-1}X\beta + (I - \rho W)^{-1}\epsilon. \end{aligned} \quad (2.13)$$

Grâce à certaines hypothèses de régularité sur W ou sur les possibles valeurs de ρ , on déduit par le théorème de l'inverse de Neumann que

$$(I - \rho W)^{-1} = I + \rho W + (\rho W)^2 + (\rho W)^3 + \dots .$$

En combinant la dernière équation avec (2.13), on trouve que

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

où $\mu = (I - \rho W)^{-1}\epsilon$. Selon LeSage et Pace [55], I donne l'effet direct, l'effet sur la variable

¹⁶. Sa méthode est qualifiée de grossière par Nwaogu. Il est l'avis de l'auteur du présent ouvrage que l'interprétation de LeSage et Pace est adéquat.

dépendante d'une variation d'une des variables spécifiques au pays. Le terme ρW révèle l'effet indirect sur les pays voisins que l'on nomme les retombées.

L'interprétation du modèle SAR et l'investissement Les motivations pour investir des firmes sont variées. Suivant [15], nous considérons le modèle SAR suivant

$$\ln(FDI) = \alpha_0 + \alpha_1 \text{Host variables} + \alpha_2 \text{Surounded-Market Potential} + \rho W \ln(FDI) + \epsilon,$$

o Host variables représentent les variables liées au pays hôte de l'investissement et Surounded-Market Potential, la variable mesure le potentiel des marchés voisins du pays hôte. Ainsi, il est possible de relier les types d'investissement avec les différentes combinaisons de signes de ρ et α_2 .

Le premier type d'investissement est l'*investissement horizontal*. La motivation d'un tel type d'investissement est l'accès aux marchés en évitant les frictions d'échange comme les coûts de transport ou des politiques protectionnistes par exemple. Dans sa plus simple forme, la seule motivation est la proximité du marché du pays hôte afin de réduire les coûts de transport. Dans un tel cas, aucun effet spatial n'est observé donc $\rho = 0$ mais de plus, aucun autre marché n'est visé d'où $\alpha_2 = 0$. Cette situation peut avoir lieu lorsque les mesures protectionnistes des pays voisins de celui visé par l'investissement sont très élevés. Dans le cas contraire, i.e. les mesures protectionnistes sont faible comparativement aux effets de friction, alors l'investissement dans un pays peut être considéré comme une plateforme d'exportation vers les marchés se situant dans les régions voisines. Puisque la firme choisit un emplacement privilégié pour son investissement, on s'attend donc à un ρ négatif reflétant le fait que la firme substitue l'investissement des pays voisins à celui du pays choisi. Il est certain que si la firme optimise l'emplacement de son investissement, l'importance du marché du pays hôte de l'investissement est donc corrélée avec l'investissement proposant ainsi un α_2 positif.

Nous ne pouvons que citer l'explication limpide que donnent Blongien et coll. dans [15] sur l'*investissement vertical*. " Dans sa forme la plus simple, l'investissement vertical est un

modèle d'investissement où les firmes multinationales évaluent tous les marchés potentiels afin de trouver celui qui leur procure le plus petit coût de l'activité qu'elles souhaitent relocaliser." [15]. Le ρ estimé est négatif encore une fois dû à l'effet de substitution précédemment mentionnée. La différence entre le modèle d'investissement vertical et horizontal réside dans le fait que le choix fait par la firme dans le modèle vertical est exclusivement basé sur le coût de relocalisation et ne prend pas en considération les marchés voisins. Par conséquent, un coefficient non significatif est associé à la variable Surrounded-Market Potential.

Finalement, [6] introduit une variante du modèle d'investissement vertical appelé *investissement direct étranger vertical complexe*. Dans ce modèle les firmes dispersent leurs activités de production dans différentes localisations. Ainsi, la proximité de fournisseurs, possiblement situés dans des régions voisines à celle où l'investissement est effectué, a un effet positif sur l'investissement. Le modèle permet également de tenir compte de la présence de ressources non mobiles telles les ressources naturelles encore une fois situées dans le pays hôte ou les pays voisins. Il va sans dire que les firmes qui s'installent dans un pays n'ayant pas les ressources ont fait face à un choix où la présence d'infrastructures, d'institutions bien établies, de mains-d'oeuvre qualifiées ou d'autres facteurs ont contrebalancés les frictions comme les coûts de transport ou les tarifs douaniers. Ce modèle d'investissement se traduit économétriquement par un signe positif de ρ . Ici encore, la proximité des autres marchés n'est pas considéré impliquant que α_2 est non-significatif.

Le Tableau 2.2, inspiré de [15, Table 1], résume quels sont les signes des paramètres à estimer selon le modèle d'investissement. Il est important de noter que ces relations se basent sur le comportement de firme et que l'utilisation de données agrégées ne peut que refléter l'effet global dominant net.

Méthode d'estimation La section 2.3.2 expose les défaillances des moindres carrés lorsque la présence de dépendance spatiale est détectée. Ceci constitue l'un des principaux inconvénients du modèle SAR. De plus, lorsqu'il est utilisé avec des données de panel, l'aspect

Modèle d'investissement	Signe de $\hat{\rho}$	Signe de α_2
Horizontal pure	0	0
Plateforme d'exportation	-	+
Vertical	-	0
Vertical complexe	+	0/+

Tableau 2.2 – Signe des différents paramètres du modèle SAR d'investissement selon le modèle d'investissement

dynamique pouvant être ajouté au modèle rend cet inconvénient encore plus litigieux.¹⁷ Les deux méthodes les plus répandues sont le maximum de vraisemblances et les moments généralisés. Nous illustrons ici comment performant ces deux techniques en nous penchant sur quelques articles empiriques.

La méthode des moments généralisés est utilisée par Baltagi et coll. [6]. Ceux-ci utilisent des données sur l'investissement direct de firmes américaines dans 51 pays, excluant les États-Unis, entre 1989 et 1999 pour étudier la pertinence de l'effet du 3^e pays. En débutant avec des régressions ne tenant pas compte de la dimension spatiale, les auteurs évaluent successivement des modèles où la dimension spatiale est incluse sur des données agrégées et puis sur les données des firmes du secteur manufacturier. L'un des résultats de cet article est que l'aspect spatial considéré est significatif, point qui donna lieu à de multiples études empiriques sur l'effet de 3^e pays. Cette significativité à un niveau de 99% n'est observable que lorsque l'investissement est mesuré à l'aide des ventes des firmes, l'autre mesure de l'investissement étant le stock de capital. De plus, une analyse sur la robustesse de leurs résultats est effectuée en utilisant différente définition de poids pour la matrice W . Il est à noter que ces tests donnent des résultats qualitativement inchangés, mais dont la significativité, elle, est affectée passant de 99% à 90%.

Dans sa thèse de doctorat [41], Hong opte également pour la méthode des moments généralisés. Il propose d'étudier l'investissement direct en Chine, qui est l'un des sujets les plus traités dans la littérature empirique, en utilisant un modèle SAR dynamique qui

¹⁷. Voir [52] pour une exposition complète des techniques d'estimation économétrique spatiale avec des données de panel. Le sujet y est traité de façon théorique très rigoureusement.

évidemment utilise des données de panel. L'aspect dynamique de sa démarche le distingue des autres études empiriques. Le modèle SAR dynamique qu'il utilise est le suivant.

$$\begin{aligned} \ln(FDI_{it}) = & \alpha_0 + \alpha_1 \text{Host variables} + \alpha_2 \text{Surrounded-Market Potential} \\ & + \rho W \ln(FDI) + \gamma \ln(FDI_{i(t-1)}) + \epsilon. \end{aligned} \quad (2.14)$$

Il est à noter que l'équation (2.14) ne tient pas compte des caractéristiques du pays investisseur. Ce choix est justifié par Hong en régressant séparément l'équation (2.14) pour chaque pays investisseur. Ses résultats proposent également que l'effet de 3^e pays est non négligeable et que son omission cause des biais considérables. L'ajout du terme $\gamma \ln(FDI_{i(t-1)})$ ajoutant l'aspect dynamique à son analyse n'affecte en rien les conclusions qualitatives, mais atténue l'importante de l'effet d'agglomération pour certain pays, i.e. la convergence des investissements dans des régions précises. Ceci laisse à penser que l'investissement est donc un processus dynamique qui, si négligé, produit des inférences potentiellement erronées.

La deuxième méthode d'estimation la plus utilisée est celle du maximum de vraisemblance. L'un des articles pionnier sur l'effet du 3^e pays est [15] où cette technique est mise à profit. Tout comme [6], Blongien et coll. étudient comment l'omission de l'effet du 3^e pays dans le cadre d'un modèle bilatéral d'investissement affecte les estimés. L'article se démarque par son analyse rigoureuse sur la question de la robustesse des résultats face à la spécification des modèles, mais également des échantillons. Pour ce faire, Blongien et coll. ont à leur disposition plusieurs échantillons allant de 1983 à 1998 de l'investissement des États-Unis dans plus de 35 pays. La mesure de l'investissement utilisée est la valeur des ventes des entreprises américaines basées à l'étranger. Leur spécification est donné par

$$\ln(FDI) = \alpha_0 + \alpha_1 \text{Host variables} + \alpha_2 \text{Surrounded-Market Potential} + \rho W \ln(FDI) + \epsilon.$$

où la variable Surrounded-Market Potential est définie par

$$\text{Surrounded-Market Potential}_i = \sum_{j \neq i} \frac{1}{d_{ij}} PIB_j.$$

L'étude de la robustesse des résultats par rapport à l'échantillonnage est faite avec la comparaison entre les pays membres de l'ODCE et les non-membres. La motivation de cette comparaison provient de [14] et [16] où des différences substantielles sont trouvés avec différents échantillons pour le modèle économétrique de gravité. Fait marquant, Blongien et coll. obtiennent que les effets fixes de périodes suffisent à contrôler les effets spatiaux dans une analyse transversale des pays développés. Les auteurs proposent l'hypothèse que les tarifs douaniers peuvent expliquer ce phénomène.¹⁸ Les mêmes régressions sont faites sur un échantillon comptant exclusivement les pays européens membres l'OCDE de leur échantillon afin de vérifier cette hypothèse. Les résultats produits sont surprenants, car le terme $\rho W \ln(FDI)$ n'est plus significatif lorsqu'ils utilisent des données agrégées. Ainsi, ils poussent plus loin leurs analyses en utilisant des données au niveau des firmes et constatent que 10 des 11 secteurs étudiés considèrent l'effet du 3^e pays comme significatif.

Panel non-balancé L'estimation des modèles spatiaux est une branche de recherche très active en économétrie et a maintenant atteint un bon niveau de maturité comme en témoigne l'ouvrage de Lee et Yu [52] contenant plus de 160 pages entièrement consacré à ce sujet. Comme le mentionne Wei dans sa thèse de doctorat [71], plusieurs méthodes d'estimation ont été développées [48] [49] et [50] pour des panels de données balancées. Nous traitons ici exclusivement d'un seul article proposant une méthode d'estimation du modèle SAR avec un panel non balancé.

Plusieurs méthodes existent pour aborder le problème d'un panel non balancé, par exemple l'exclusion des données, des algorithmes de maximisation de l'espérance (algorithme EM) ou encore l'imputation multiple. Cependant, chacune de ces méthodes comporte un risque important dans le cadre de panel spatial. En effet, l'exclusion d'une localisation spatiale et des données s'y rattachant du point de vue mathématique peut détruire la structure originale du modèle SAR en modifiant la matrice W . LeSage et Pace [54] utilisent les algorithmes EM avec succès, mais comme le mentionnent Lee et Liu [50] l'estimateur obtenu par l'utilisation conjointe de la méthode du maximum de vraisemblance peut rencontrer des

18. Il est vrai que Blongien et coll. présente cette hypothèse. Cependant, elle a préalablement été suggérée par un arbitre.

problèmes lorsque la dépendance spatiale est d'ordre supérieur à 1.

Wei dans [71] propose d'utiliser d'astucieuse manière la méthode des moments généralisée afin de traiter le cas où des observations de variables dépendantes sont manquantes. Sa méthode se résume ainsi " ...déterminer l'espérance des données manquantes comme fonction des variables indépendantes observées et des paramètres estimés puis remplacer les données manquantes." [71]. Nous présentons ici le moment clé de sa méthode. Le modèle SAR utilisé est donné par

$$Y_n = \lambda_0 W_n Y_n + X_n \beta_0 + \epsilon_n, \quad \epsilon \sim N(0, \sigma^2). \quad (2.15)$$

Posons

$$Y = \begin{pmatrix} Y_n^o \\ Y_n^* \end{pmatrix}$$

où Y_n^o sont les données observés et Y_n^* les données manquantes. Une réécriture habituelle de l'équation (2.15) est

$$(I - \lambda_0 W_n) Y_n = X_n \beta_0 + \epsilon_n.$$

Soit J_n une matrice telle que $J_n Y_n = Y_n^*$, alors en utilise l'équation précédente, on a

$$Y_n^* = J_n Y_n = J_n (I - \lambda_0 W_n)^{-1} X_n \beta_0 + J_n (I - \lambda_0 W_n)^{-1} \epsilon_n.$$

Donc,

$$E(Y_n^*) = E(J_n (I - \lambda_0 W_n)^{-1} X_n \beta_0).$$

Ainsi, le moment utilisé par Wei est donné par

$$E(U_n) = \epsilon_n - (I - \lambda_0) \begin{pmatrix} 0 \\ J_n (I - \lambda_0 W_n)^{-1} \epsilon_n \end{pmatrix} = 0.$$

L'estimateur obtenu par Wei est consistant et sans biais moyennant l'ajout de quelques hypothèses de régularité. Ce qui est exceptionnel de sa méthode est que Wei détermine explicitement la matrice de poids optimale pour sa technique.

2.3.3 Modèle d'erreur spatiale

Une autre façon pour la dépendance spatiale de se manifester est par le terme d'erreur. Cette situation est possible par l'introduction d'erreurs de mesure dépendantes de la localisation tel qu'illustré dans l'exemple du taux de chômage mesuré spatialement par le lieu de résidence. Dans cette section, nous présentons le modèle d'erreur spatiale qui tient compte de cette particularité de la dépendance spatiale ainsi que le test I de Moran permettant de déceler l'autocorrélation linéaire spatiale du terme d'erreur. Cependant, l'espace consacré à ce modèle reflète sa faible popularité dans les travaux empiriques. Ainsi, nous nous attardons qu'à un seul article empirique utilisant ce modèle qui traite également du problème des données de panel non balancé.

Le modèle d'erreur spatiale est donné par

$$\begin{aligned} y &= X\beta + \mu & (2.16) \\ \mu &= \lambda W\mu + \epsilon, \quad \epsilon \sim N(0, \sigma^2). \end{aligned}$$

On constate tout de suite que la dimension spatiale est exclusivement contenue dans le terme d'erreur μ . La forme particulière de ce modèle illustre malgré tout comment la dimension spatiale peut être omise facilement puisque cette dernière est contenue dans le terme d'erreur. Si aucun test n'est fait pour vérifier la présence d'autocorrélation spatiale, le modèle estimé est

$$y = X\beta + \mu \tag{2.17}$$

$$\tag{2.18}$$

où le terme d'erreur ne satisfait pas les hypothèses du théorème de Gauss-Markov. Ainsi,

l'estimateur des moindres carrés obtenu par l'estimation de ce modèle est donc biaisé et même non consistant.

Test I de Moran

Si la spécification du modèle ne prévoit pas de dépendance spatiale pour la variable dépendante et que nous utilisons des données contenant une dimension spatiale, il est probable que celle-ci soit contenue dans les résidus, i.e. qu'il y a présence d'autocorrélation spatiale du terme d'erreur. Pour traiter ce sujet, nous suivons de très près le livre de LeSage [53]. Un test répandu pour vérifier la présence d'autocorrélation spatiale dans le terme d'erreur est le Test I de Moran. La statistique I de Moran est donnée par

$$I = \frac{e'W e}{e'e}$$

où e est le vecteur des résidus obtenus de l'estimation des moindres carrés de l'équation (2.17) et W la matrice des poids spatiaux. Les deux caractéristiques importantes de cette statistique sont

$$\begin{aligned} E(I) &= \frac{\text{Tr}((Id - X(X'X)^{-1}X')W)}{n - k} \\ \text{Var}(I) &= \frac{\text{Tr}(MWMW') + \text{Tr}(MW)^2 + (\text{TR}(MW))^2}{(n - k)(n - k + 2)} \end{aligned}$$

où l'opérateur Tr dénote la trace, X est la matrice des variables explicatives de dimension $n \times k$ et $M = (Id - X(X'X)^{-1}X')$ pour simplifier la notation. Les détails des calculs sont donnés à la page 102 de [4]. Il est possible de montrer qu'asymptotiquement cette statistique suit une loi normale centrée réduite. Ainsi, la statistique à tester est la suivante.

$$I^* = \frac{I - E(I)}{\sqrt{\text{Var}(I)}} \sim N(0, 1)$$

Asselin et Florax [5] effectuent des simulations sur de petits échantillons pour déceler l'autocorrélation spatiale du terme d'erreur en utilisant le test I de Moran. Ils obtiennent des résultats assez satisfaisants pour affirmer que cette statistique peut également être utilisée

sur de petit échantillon. L'une des hypothèses fondamentales de ce test est que toutes les autres formes tendances sont éliminées avant que le test soit fait.¹⁹

2.3.4 Panel non-balancé

L'impopularité empirique du modèle d'erreur spatiale est due au fait que la dépendance spatiale est entièrement contenue dans le terme d'erreur rendant pratiquement impossible l'évaluation de son impact sur les résultats et donc produit une analyse incomplète de la situation. Ainsi, peu d'articles l'utilisent. L'un des rares à le faire est [32] que nous résumons ici.

Egger et coll. [32] étudient la compétition dans la course à la réduction des taxes entre les états des États-Unis de 1975 à 1999. Le modèle économétrique préconisé par les auteurs est un modèle mixte, i.e. un modèle SAR dont le terme d'erreur tient compte de l'autocorrélation spatiale. Ce modèle appelé modèle de Kelejian-Prucha [44] s'écrit comme suit.

$$\begin{aligned}\tau_{it} &= \rho_1(W_{N_t}\tau)_i + X_{it}\beta + \lambda_t + \mu_{it} \\ \mu &= \rho_2 W_{N_t}\mu + \epsilon \\ \epsilon_{it} &= \mu_i + \nu_{it}\end{aligned}$$

où l'indice t dénote la période, le i l'état. De plus, τ est utilisé comme variable du taux de taxation, X un vecteur de variables explicatives et W_{N_t} la matrice des poids indexés par N_t . La raison de l'indice N_t est que les auteurs utilisent des données de panel non balancé où N_t représente le nombre de données utilisé à la période t .²⁰ Les constantes λ_t et μ_i sont insérés afin de prendre en considération les effets fixes de périodes et de régions respectivement. Leur technique d'estimation découle de l'enchaînement de plusieurs estimations. En un premier temps, un estimateur consistant pour ρ_1 et β est obtenu avec la méthode en 2 étapes de moindres carrés (2SLS) où la variable instrumentale est un terme de dépendance spatiale de τ d'ordre supérieur ou égal à 2. La deuxième étape consiste à utiliser un estimateur des

19. Voir [13] pour plus de détails sur ce sujet.

20. Nous utilisons ici la même notation que [32]. Nous saluons ici cette notation, car une notation plus compacte incorporant le terme des effets fixes dans la première équation est possible. Cependant, de ce fait, la méthode d'estimation utilisée devient plus nébuleuse à comprendre.

moments généralisés pour estimer ρ_2 ainsi que les effets fixes de temps et de périodes telles que proposées par [43]. Finalement, en utilisant une transformation de type Cochran-Orcutt, un estimateur tenant compte des effets aléatoires est obtenu par la méthode généralisée des moindres carrées. L'originalité de cet article est indéniablement la spécification de son modèle ainsi que la technique d'estimation. De plus, il est important de noter que leurs estimés sont consistents et suivent asymptotiquement une loi normale [43].

Chapitre 3

Données

Dans ce chapitre, nous décrivons toutes les variables nécessaires à notre étude et plus particulièrement l'investissement direct étranger américain au Canada. Nous abordons en un premier temps les raisons qui nous poussent à restreindre notre étude exclusivement aux investissements provenant des États-Unis. Puisque les sources de nos données sont très variées, nous prenons soin de les rattacher à chacune de nos variables. Finalement, nous traitons des mesures de dispersion de nos données en insistant sur la variabilité des variables. Ce point étant crucial pour la sélection de la méthode d'estimation de panel que nous aborderons au Chapitre 4.

3.1 L'investissement direct américain au Canada

Notre étude se base sur un panel de données de l'investissement direct américain de 2004 à 2011 au Canada dans chacune des provinces canadiennes à l'exception de l'Île-du-Prince-Édouard. L'investissement direct étranger au Canada se distingue de celui des autres membres de l'OCDE de par sa situation géographique. Le Canada partage ses frontières terrestres avec un seul pays, les États-Unis. Ainsi, sa situation géographique lui confère un accès privilégié à une des plus importantes économies du monde. Cette proximité se reflète dans l'investissement direct étranger provenant des États-Unis. Par exemple en 2011, plus

de 57% de l'investissement direct étranger provenait des États-Unis.¹ D'autres pays tels que l'Allemagne, la France et le Japon sont également des investisseurs étrangers importants, mais de moins grande envergure pour l'économie canadienne.

Nous limitons notre étude aux investissements directs étrangers américains pour 2 raisons. Premièrement, notre voisin américain est de loin le plus important pays investisseur au Canada. Le Tableau 3.1 montre que la moyenne annuelle du pourcentage que représente l'investissement direct américain de tout l'investissement direct étranger est de plus de 55% pour 9 provinces sur 10. Ceci démontre clairement l'importance de l'investissement direct des États-Unis pour les économies provinciales canadiennes.

Provinces	Pourcentage (%)
Alberta	69,37
Colombie Britannique	69,11
Île -du-Prince-Édouard	70,12
Maniotba	65,69
Nouveau Brunswick	55,14
Nouvele écosse	70,07
Ontario	62,64
Québec	56,23
Saskatchewan	56,00
Terre-Neuve-Laberador	37,14

Tableau 3.1 – Moyenne annuelle du pourcentage de l'investissement direct américain de l'investissement direct étranger de 2004 à 2011

Une autre motivation pour restreindre l'étude de l'investissement direct étranger à celui des États-Unis réside dans un manque de données disponibles. En effet, pour des pays comme l'Allemagne ou l'Angleterre le pourcentage de données manquantes atteint jusqu'à 50% pour certaines provinces. La manière usuelle de contourner ce problème est tout simplement l'exclusion de la province dont certaines des données sont manquantes. Cependant, une telle approche entraînerait l'exclusion de toutes les provinces. Cette problématique connaît récemment une grande popularité dans la littérature empirique spatiale et tout particulièrement dans le cas qui nous préoccupe ici c'est-à-dire des observations manquantes de la

1. Pourcentage obtenu à partir des données de [23].

variable dépendante [71, chapitre 1], [72], [54].

L'exclusion de l'Île-du-Prince-Édouard de notre étude est justifiée par le fait que l'investissement direct américain dans cette province en 2005 est non disponible. Il faut comprendre que dans le contexte de possible dépendance spatiale l'omission ou l'exclusion d'une unité géographique signifie que la structure spatiale étudiée est différente de la structure réelle. Nous répondons à cette critique très populaire dans la littérature par deux arguments. Premièrement, nous avons effectué des tests d'autocorrélation spatiale utilisant le multiplicateur de Lagrange [8]; un sans l'Île-du-Prince-Édouard et un autre avec une imputation multiple de la donnée manquante. Dans les deux scénarios, nous observons une autocorrélation spatiale d'amplitude similaire.² Une autre option est proposée par LeSage et Pace [54] préconisant l'utilisation du maximum de vraisemblance conjointement avec des algorithmes EM pour produire des imputations multiples. Des estimations efficaces sont produites avec leurs techniques, mais sous l'hypothèse que les erreurs soient normalement distribuées. Nous prenons soin ici de dire qu'il ne suffit pas qu'elles le soient pour tout l'échantillon, mais également pour chaque unité géographique sans quoi l'estimation par maximum de vraisemblance avec des données de panel spatial est inefficace. Les graphiques D.1 à D.9 nous montrent une estimation par noyau de densité de la fonction de densité des résidus de nos estimations³ ainsi que la distribution normale associée. On y voit clairement que les résidus ne suivent pas une loi normale pour chaque province écartant ainsi l'approche de LeSage et Pace.⁴

3.2 Description

Étant donné la grande diversité des sources ainsi que des unités de mesure des données que nous utilisons, nous croyons qu'il est essentiel de bien les identifier. De plus, nous décrivons en détail comment les variables servant à mesurer les marchés voisins ont été construites et à partir des quelles données.

2. Les résultats des tests utilisant les imputations multiples peuvent être obtenus sur demande à l'auteur.

3. Les résidus sont ceux du modèle de base donné par l'équation 4.2 et sont obtenus en utilisant la méthode proposée par Drukker, Prucha, et Raciborski dans [31].

4. Cette observation concorde également avec l'estimation par noyau de densité des fonctions de densité de l'investissement direct américain faites pour chacune des provinces ainsi qu'avec des tests de normalité de Shapiro-Wilk. Les graphiques et résultats peuvent être obtenus sur demande à l'auteur.

Nous débutons par l'investissement direct américain. Les données sont tirées de [19], [20] [21], [22] et [23]. Ces catalogues de Statistique Canada publiés annuellement depuis 2004 contiennent des données agrégées sur les dépenses en investissement en construction non résidentielle, en outillage et en machinerie, et ce en million de dollars courants. Chaque numéro donne de l'information sur 3 années consécutives où les investissements de la première année sont ceux qui ont été réellement effectués, ceux de la deuxième période sont basés sur des estimations préliminaires de Statistique Canada et des prévisions sont fournies pour la dernière année. Nous considérons donc de 2004 à 2009 les investissements qui ont réellement été faits durant cette période. Les valeurs de 2010 et 2011 proviennent de [23] et sont donc respectivement des estimations préliminaires et des prévisions. Nous notons que nous avons choisi de considérer les prévisions de 2011 étant donné la courte période d'étude de notre panel.

Il est certain que l'investissement direct étranger dépend de plusieurs facteurs. Cependant, nous considérons ceux qui sont le plus fréquemment utilisés dans les récentes études empiriques par exemple [6] [15] et [14]. La variable *Coût_unitaire*, qui donne le coût moyen d'une heure de main-d'oeuvre, est mesurée en dollars par unité de *PIB* réel pour chacune des provinces et est tiré du tableau CANSIM 383-0011.⁵ Aucune autre transformation n'a été apportée à cette variable. Le produit intérieur brut, *PIB*, de chacune des provinces provient du tableau CANSIM 379-0025. Le *PIB* est donné en prix de base en million de dollars canadiens. Nous considérons le logarithme naturel du *PIB*, $\ln(PIB)$, dans nos régressions. Afin de mesurer la qualité de la main-d'oeuvre, nous utilisons 2 proxis soit *NB_etudiant* qui donne le nombre d'étudiants qui sont dans tous les types d'institutions postsecondaires publiques. Ces données proviennent du tableau CANSIM 477-0019.⁶ La deuxième proxi, *NB_diplome* est le nombre de diplômés postsecondaires de la population active. La population active telle que définie par Statistique Canada est constituée de toute personne de 15 ans ou plus exerçant un emploi ou exploitant une entreprise ou étant chômeur. Ces données

5. Ce tableau est maintenant remplacé par le tableau 383-0029

6. Il est important de noter que le nombre d'étudiants de l'Université de Regina ne sont pas considéré de 2005 à 2008 puisque celui-ci n'est pas disponible.

constituent des estimations basées sur l'Enquête de la population active qui sont tirées du tableau CANSIM 282-0003. Nous transformons ces deux variables de sorte que leurs unités de mesure soit par 100 000 individus. Par la suite, le logarithme naturel est appliqué lors des régressions. Nous considérons également la population, donnée en million, provenant du tableau CANSIM 051-001. Cette variable comprend tous les groupes d'âge. Nous mentionnons que la population âgée entre 15 et 64 ans a été également considérée. Puisque les résultats obtenus ne diffèrent pas de ceux présentés ici, nous ne les avons pas inclus.⁷ L'impact de la taxation des entreprises est mesuré à l'aide de la variable *METR* qui représente le taux marginal effectif de taxation agrégé des corporations canadiennes en pourcentage. Nous utilisons exclusivement [28] comme source pour cette variable.⁸ Nous suggérons aux lecteurs [25] pour obtenir un résumé de la méthodologie concernant le calcul du taux marginal effectif de taxation.

Les variables *sur_market_c*, *sur_only_usa* et *sur_can_usa* mesurent le potentiel des marchés voisins de chacune de province. Le marché potentiel exclusivement canadien est mesuré par *sur_market_c*, celui des états américains⁹ possédant une frontière commune avec les Canada par *sur_only_usa* et celui du marché canadien et des états américains limitrophes simultanément par *sur_can_usa*. Ces variables sont une moyenne pondérée des PIB des provinces canadiennes et des états considérés et se mesure donc en *million\$/km*. Nous définissons *sur_market_c_{i,t}* pour une province *i* au temps *t* de la manière suivante.

$$sur_market_c_{i,t} := \sum_{j=1, j \neq i}^9 PIB_{j,t} \cdot \frac{1}{d_{j,i}}$$

où $d_{i,j}$ représente la distance entre les capitales des provinces *i* et *j*. Les variables *sur_only_usa* et *sur_can_usa* sont définies de manière similaire en considérant les états américains et les provinces canadiennes et états américains respectivement. La présence de $d_{i,j}$ sert à contrôler pour le coût de transport et possiblement pour d'autres frais fixes liés à la distance. Nous

7. Ces résultats sont disponibles sur demande à l'auteur.

8. Voir [26] et [27] pour d'autres sources.

9. Voir l'Annexe A pour une liste exhaustive des états américains considérés.

ne faisons pas exception et mesurons la distance avec "great circle area"¹⁰ et ce en *km*. Les *PIB* en dollars américains des états considérés proviennent du Bureau of Economic Analysis. Il est important de noter que nous n'avons pas transformé les PIB en dollars canadiens avec le taux de change pour des raisons de simplicité. Le taux change annuel moyen entre le dollar américain et canadien de 2004 à 2011 est demeuré très près de la parité et donc cette simplification n'engendre que de très faibles fluctuations qui n'influencent en rien les résultats obtenus. Finalement, la variable *sur_market_metro* mesure également le potentiel des marchés canadiens voisins, mais en utilisant les distances entre les métropoles des provinces.

Variabiles	Moyenne	Écart type	Min.	Max.	Obs.
Investissement	3 492,564	3 976,926	257,1	12 506,3	72
METR	24,592	9,674	2,9	43,9	63
Nombre d'étudiants	0,192	0,22	0,024	0,761	72
Nombre de diplômés	1 ,547	1,735	0,183	6,016	72
Coût unitaire	0,662	0,069	0,481	0,78	72
Population	3,658	3,985	0,506	13,366	72
PIB	132 028,7	146 622,6	15 763,2	496 953,9	72
Marché can.	760,935	275,217	429,475	1 426,074	72
Marché amé. et can. (combinés)	3 404,781	1 310,616	1 795,293	6 130,051	72
Marché amé.	2 643,846	1 261,738	1 301,767	5 590,203	72
Marché can. (Métropoles)	793,323	228,402	492,885	1 209,833	72

Tableau 3.2 – Statistiques descriptives

Le tableau 3.2 présente des mesures de dispersion de nos données sans tenir compte de l'aspect de panel. Ceci nous permet d'avoir une vue d'ensemble de toutes les données et de leurs variabilités. Dans un premier temps, le nombre d'observations est de 72 pour toutes les variables exceptées *METR* puisque le taux marginal effectif de taxation des entreprises n'est disponible qu'à partir de 2005. Une particularité des résultats du tableau 3.2 est que la moyenne du nombre d'étudiants et celle du nombre diplômé sont différentes. Nous rappelons ici, que la première représente la moyenne provinciale du nombre d'étudiants dans tous les types d'institutions publiques entre 2004 et 2011 alors que la moyenne de la variable

10. La formule utilisée prend en considération la forme elliptique de la terre. Ce point est parfois négligé pour des raisons de simplicité.

NB_diplome est celle de tous les diplômés postsecondaires dans la population active provinciale durant la même période couvrant une plus grande partie de la population. Cette différence se remarque également dans la grande différence de l'étendue de ces deux variables. On constate que l'étendue est assez importante pour la plupart des variables à l'exception du coût unitaire qui n'est que de 0,299 dollar par unité de PIB réel. Nous nous attardons plus précisément sur cette statistique puisqu'elle a un grand impact sur la décision de considérer uniquement les effets fixes par provinces. Les tableaux C.1 à C.9 contiennent également les statistiques descriptives des variables utilisées des provinces tenant ainsi compte de la dimension longitudinale des données. Avec cet exercice, nous constatons que l'investissement américain en Alberta passe de 7 milliards à 12 milliards (tableau C.1), alors qu'il n'augmente que de 400 millions en Nouvelle-Écosse (tableau C.5). On remarque également que le nombre d'étudiants et le nombre de diplômés postsecondaires de la population active ne varient pratiquement pas comme le démontre les statistiques du Nouveau-Brunswick avec un minimum de 0,031 et 0,0286 pour le nombre d'étudiants et de diplômés respectivement et des maximums des 0,032 et 0,303 respectivement. Une situation similaire est observée pour la Nouvelle-Écosse. La variabilité de la population est elle aussi pratiquement inexistante. Seul l'Ontario avec une augmentation de 1 000 000 de personne est digne de mention.¹¹

11. L'étendu observé est de 1 000 000, cependant puisque la population est en constante croissance, le minimum et maximum de cette variable est observé en 2004 et 2011 respectivement pour l'Ontario.

Chapitre 4

Spécification des modèles

Dans ce chapitre, nous présentons les spécifications économétriques spatiales que nous utilisons. Nous prenons soin de fournir pour chaque variable le signe de la corrélation attendue d'après la littérature actuelle ainsi qu'une justification économique adéquate.

4.1 Modèle de gravité standard

Afin d'examiner l'impact de la dépendance spatiale de l'investissement direct américain, noté FDI , dans les provinces canadiennes, nous débutons avec un modèle standard [14] [15] [63] donné par

$$\ln(FDI_{it}) = \alpha_0 + \alpha_1 \text{Variables Hôtes}_{it} + u_i + \epsilon_{it} \quad (4.1)$$

où l'indice i indique la province et t la période. Le terme `Variable_Hôtes` désigne les variables `Cout_unitaire`, `PIB`, `POP`, `NB_etudiant`, `NB_diplome` et `METR`. Il est à noter que les variables `POP`, `NB_etudiant` et `NB_diplome` sont sous forme logarithmique. Cette spécification log-linéaire s'inspire des précédentes études [14] [15] [6]. De plus, ce modèle s'inscrit dans le cadre de la théorie de la gravité des investissements bilatéraux. Les variables liées aux pays investisseurs sont toutes regroupées à l'intérieur du terme α_0 . Une telle approche est également utilisée dans [15].

En accord avec la littérature existante, plus le *PIB* de la province hôte est élevé plus l'investissement est élevé également. Ceci se justifie par le fait que le *PIB* sert de proxy pour mesurer le marché potentiel de la province hôte de l'investissement. Avec la spécification donnée par l'équation 4.1, le coefficient de la variable *POP* représente l'élasticité de l'investissement par rapport à la population qui sera négative. En effet, une augmentation de la population entraîne une diminution de la richesse provinciale (*PIB* par habitant) puisque que le *PIB* demeure constant lorsque l'on calcul cette élasticité. Un autre facteur important est le coût de la main-d'oeuvre qui est mesuré par la variable *Cout_unitaire*. L'investissement diminue lorsque les investisseurs font face à des coûts de main-d'oeuvre élevés proposant ainsi une corrélation négative entre l'investissement direct étranger et le coût unitaire de la main-d'oeuvre. Il est certain qu'un coût élevé de la main-d'oeuvre peut cacher une main-d'oeuvre plus qualifiée. Afin de tenir compte de cette éventualité, nous utilisons les variables *NB_etudiant* et *NB_diplome* comme proxy pour déterminer l'abondance de main-d'oeuvre qualifiée dans la province. Ainsi, pour des raisons de productivité plus élevée, mais également pour réduire les coûts d'implantation de l'investissement, l'investissement direct étranger devrait être corrélé positivement avec ces deux variables. Finalement, nous mesurons l'impact des dispositions fiscales liées à la taxation des entreprises en utilisant la variable *METR* mesurant le taux marginal effectif de taxation des corporations. A priori, les multinationales favorisent l'investissement là où les taxes sont faibles proposant ainsi une corrélation négative entres *METR* et *FDI*. Nous argumentons que la corrélation positive ne peut être écartée. En effet, un haut taux marginal effectif de taxation des entreprises peut indiquer, particulièrement pour des pays, états ou provinces développés, de meilleures infrastructures publiques dans la mesure où l'infrastructure est une variable omise qui est positivement corrélé avec le niveau de taxation. Celles-ci facilitent l'accès notamment aux ressources naturelles nécessaires, mais également au marché intérieur de la province hôte et à ses marchés voisins. De ce fait, le signe du coefficient *METR* est indéterminé. Nous mentionnons que des analyses ont été faites également avec l'utilisation d'une variable donnée par les dépenses en immobilisation dans les services publics¹. Puisque les estimés obtenus n'étaient pas statistiquement significatifs et ne changeaient en rien nos résultats, nous avons

1. Voir tableau CANSIM 029-0005 pour une description détaillée.

choisi de ne pas les présenter.²

Nous terminons cette section en mentionnant que nous ne considérons que les effets fixes par provinces. Cette approche, qui est utilisée également dans les spécifications subséquentes, est justifiée premièrement par la variabilité temporelle presque inexistante des variables considérées telles que mentionnées à la section 3.2. De plus, les effets fixes sont en mesure de capter, par exemple, l'abondance de ressources naturelles ou encore de main-d'oeuvre, qui ne varie pas dans le temps. De façon plus général, les facteurs qui s'avèrent pertinents à l'étude de l'investissement direct étranger varient très peu au cours de la période d'étude, ainsi l'inclusion des effets fixes provinciaux permet de réduire le nombre de variables explicatives et de conserver seulement celles d'intérêt à notre étude. Ceci semble être en contradiction à première vue avec les événements économiques de cette période, c'est-à-dire la récession débutant en 2008 et se terminant en 2009. Nous avons effectué des estimations avec des variables binaires pour ces deux années. Aucune d'entre elles n'était statistiquement significative, nous les avons donc écartées.

4.2 Spécifications spatiales

Nous avons modifié l'équation 4.1 de la manière suivante afin de modéliser l'interdépendance spatiale des données. Pour la province i à la période t , nous estimons

$$\begin{aligned} \ln(FDI_{it}) &= \alpha_1 \text{Variables Hôtes}_{it} + \text{Sur_market Potential} & (4.2) \\ &+ \rho \sum_{j=1}^9 w_{ij,t} \ln(FDI_{j,t}) + u_i + \nu_{i,t} \\ \nu_{i,t} &= \lambda \sum_{j=1}^9 w_{ij,t} \epsilon_{j,t} + \epsilon_{i,t} \end{aligned}$$

où $w_{ii,t} = 0$ pour tout i et $\epsilon_{j,t}$ représente le terme d'erreur stochastique de la province j au temps t . Cette formulation est plus générale que ce qui est nécessaire puisque les poids $w_{ij,t}$

2. Ces résultats peuvent être obtenus sur demande à l'auteur.

ne varient pas dans le temps dans la présente étude.³ Le poids $w_{ij,t}$ sont les composantes de la matrice de poids spatiaux W telle que définie dans la section 2.3.1. Nous mentionnons que la forme de W suit celle qui est la plus largement utilisée dans la littérature sans autre considération. Ainsi, la structure spatiale utilisée ici est donc exogène et constitue une hypothèse importante de notre étude. Il est à noter que la distance minimale de 280,91 km utilisée est celle entre les villes d’Halifax en Nouvelle-Écosse et de Fredericton au Nouveau-Brunswick.

Nous incluons en premier lieu *Cost_unitaire* et *PIB* comme variables hôtes pour constituer notre modèle de base pour les estimations spatiales étant donné que nous considérons ces deux variables comme des facteurs clés de l’investissement direct américain. Nous ajoutons par la suite séparément les autres variables de contrôle précédemment mentionné. Cette façon de procéder est utilisée dans le but de bien saisir l’impact de leur introduction dans les estimations et de tester la robustesse des résultats face à l’interdépendance spatiale. Finalement, une spécification utilisant toutes les variables hôtes à l’exception du nombre de diplômés est considérée.

Le terme $\rho \sum_{j=1}^9 w_{ij,t} \ln(FDI_{j,t})$ est introduit afin de refléter l’interdépendance spatiale. Le coefficient ρ mesure son importance et son effet, soit positif ou négatif, sur l’investissement direct étranger américain et ce en capturant la proximité de l’investissement observé dans la province hôte par rapport aux autres provinces [15]. En aucun cas, on ne doit confondre ce terme avec les mesures de distances utilisées dans un modèle de gravité d’investissement bilatéral. Nous avons déjà abordé à la section 2.3.2 comment relier différentes combinaisons des signes de ρ et α_2 avec les modèles d’investissement. Une précision concernant le modèle complexe vertical doit être faite. Il est vrai que si la variable *Sur_market Potential* capture exclusivement l’effet de la demande des marchés voisins, le coefficient ne sera pas statisti-

3. La formulation la plus couramment utilisée dans la littérature est sous forme vectorielle et est donnée par

$$\begin{aligned} \ln(FDI_t) &= \alpha_1 \text{Variables Hôtes}_t + \text{Sur_market Potential} + \rho W \ln(FDI_t) + u + \nu_t \\ \nu_t &= \lambda W \epsilon_t + \epsilon_t \end{aligned}$$

où W représente la matrice de poids spatiaux invariant dans le temps défini à la section 2.3.1, u le vecteur des effets fixes par provinces et ϵ_t le vecteur des termes d’erreur.

quement significatif. Cependant, si cette variable capte également les effets d'agglomération, ce qui est plus probable lorsque les données sont agrégées, alors une corrélation positive sera observée entre l'investissement et le potentiel des marchés voisins [32].

Pour mieux tester quel marché, canadien ou américain, est visé par l'investissement direct américain au Canada, nous utilisons différentes spécifications en changeant la variable *Sur_market Potential*. Il est certain que si l'investissement suit un modèle horizontal pur ou vertical pur, il n'y aura aucun effet visible entre ces différentes spécifications. Dans le cas où le modèle est une plateforme d'exportation, il est possible de tester si le marché américain a de l'importance. C'est ce que nous faisons en ayant une spécification utilisant exclusivement le marché canadien avec la variable *sur_market_can*, exclusivement le marché américain en utilisant *sur_only_usa* et les deux marchés conjointement, c'est-à-dire dans une seule variable, avec *sur_can_usa*. Si le marché canadien est le seul visé alors le coefficient de *sur_only_usa* ne sera pas statistiquement significatif, alors que celui de *sur_market_can* et *sur_can_usa* devrait l'être. Dans le cas où seul le marché américain est visé, les coefficients de *sur_only_usa* et *sur_can_usa* seront statistiquement significatifs. La dernière possibilité est que les deux marchés soient visés. C'est pour cette raison que nous estimons également le modèle de base avec la variable *sur_market_can* et *sur_only_usa* afin de déterminer l'importance de chacun pour les investisseurs.⁴

Une particularité du modèle utilisé dans cette étude est la présence de la dépendance spatiale dans le terme d'erreur exprimée à l'aide du terme $\lambda \sum_{j=1}^9 w_{ij,t} \epsilon_{j,t}$. Bien que plusieurs ouvrages de la littérature ne l'inclut pas [15] [6], nous justifions son utilisation avec des tests statistiques présentés à la section 5.1, mais également par une conjecture d'évènements économiques. Bien que nous ayons écarté l'utilisation des effets fixes de périodes, il n'en reste pas moins que la récession de 2008-2009 est un évènement marquant pour l'économie mondiale qui a affecté chacune des provinces différemment. Durant celle-ci les gouvernements provinciaux ont mis sur pied des programmes spéciaux et temporaires afin de stimuler l'éco-

4. L'auteur préconise une approche similaire pour étudier l'investissement direct étranger autre que celui américain. Puisque le Canada et les États-Unis ont un accord de libre-échange, l'investissement direct étranger au Canada peut suivre un modèle de plateforme d'exportation visant le marché américain pour profiter de cet accord de libre-échange.

nomie. Ce genre d'initiative est modélisé par notre modèle à l'intérieur du terme d'erreur. Or puisque les provinces n'ont pas adopté des mesures de concert, leurs effets "aléatoires" sont cependant interreliés favorisant ainsi une autocorrélation spatiale du terme d'erreur justifiant ainsi la forme du modèle utilisé.

Chapitre 5

Résultats empiriques

Nous présentons dans ce chapitre les résultats de nos régressions en débutant avec ceux concernant notre choix du modèle spatiale *SAC* parmi les autres modèles de régressions spatiales évoqués dans la littérature. Nous passons ensuite à l'étude de l'ajout des variables de contrôle dans deux scénarios, c'est à dire l'inclusion ou non du taux marginal effectif de taxation des entreprises. Nous terminons ce chapitre en analysant les résultats concernant l'importance de la proximité du marché américain pour l'investissement direct américain.

Avant de continuer, nous précisons que toutes les estimations des spécifications spatiales ont subi la transformation proposée par Lee et Yu [51] puisque celle-ci permet d'obtenir des estimés efficaces comparativement à la transformation usuelle.

5.1 Sélection du modèle

Nous avons déjà abordé dans les sections 2.3.2 et 2.3.3 les modèles *SAR* et *SEM* respectivement. De plus, le modèle présenté dans l'équation 4.2 est un hybride de ces deux modèles. Cependant la littérature contient bien d'autres modèles spatiaux de panel qui représentent des candidats potentiels que nous ne devons pas négliger.

Le modèle Durbin spatial [56], [63] sous forme vectorielle est donné, pour notre cas d'étude,

par

$$\begin{aligned} \ln(FDI) &= \alpha_1 \text{Variables Hôtes} + \text{Sur_market Potential} \\ &+ \rho W \ln(FDI) + \beta W \text{Variables Hôtes} + u + \epsilon \end{aligned} \quad (5.1)$$

où W est la matrice de poids spatiales, u le vecteur des effets fixes des provinces et $\epsilon \sim N(0, \sigma_\epsilon^2)$ le vecteur d'erreur. On peut considérer ce modèle comme une généralisation du modèle *SAR*. Un autre candidat que nous avons considéré est le modèle *GSPRE* (Generalised Spatial Panel Random Effects) introduit dans [7] et donné par

$$\begin{aligned} \ln(FDI_{i,t}) &= \alpha_1 \text{Variables Hôtes}_{i,t} + u_i + \nu_{i,t} \\ u_i &= \phi \sum_{j=1}^9 w_{ij} \mu_j + \mu_i \\ \nu_{i,t} &= \lambda \sum_{j=1}^9 w_{ij} \epsilon_{j,t} + \epsilon_{i,t} \end{aligned} \quad (5.2)$$

où $w_{ii} = 0$ pour tout i , μ_i représente les effets aléatoires des provinces, et non pas des chocs aléatoires, et $\epsilon_{i,t} \sim N(0, \sigma_\epsilon^2)$ le terme d'erreur. Ce modèle se distingue par l'interdépendance spatiale entre les effets aléatoires. Ce modèle constitue un modèle potentiel de par la justification économique que nous avons fournie à la section 4.2 concernant la présence d'autocorrélation entre les termes d'erreur. Si jamais, l'autocorrélation envisagée est invariante dans le temps, ce modèle sera mieux adapté que celui présenté à la section 4.2.

Dans un premier temps, nous devons vérifier s'il y a bel et bien de l'autocorrélation spatiale dans nos données. Pour ce faire, nous utilisons la méthodologie proposée par [8]. Nous préconisons cette approche plutôt que le test de Moran étant donné que celle-ci nous permet de tester de façon robuste l'interdépendance spatiale de l'investissement, $\rho W \ln(FDI)$, face à la possibilité d'une autocorrélation spatiale entre les termes d'erreur.¹ Cette méthodologie est basée sur un test du multiplicateur de Lagrange. Nous avons effectué ces tests avec

1. Voir [8] pour plus de renseignements sur ce point ainsi qu'une description complète des tests et des statistiques utilisés.

le progiciel *splm*² disponible avec le logiciel d'analyse statistique *R*. La première ligne du

Hypothèses nulles		Lagrange Multiplier	P-value
H_0^a	Aucune dépendance spatiale Aucune autocor. spa. erreur	96,668	0,000
H_0^d	Aucune dépendance spatiale Possible existence autocor. spa. erreur	1,8679	0,062

La notation utilisée est la même que celle de [8] et [60]

Tableau 5.1 – Test d'autocorrélation spatiale du multiplicateur de Lagrange Baltagi et col. [8]

tableau 5.1 présente les résultats du test d'hypothèse avec comme hypothèse nulle l'absence d'interdépendance spatiale, $\rho = 0$ et d'effets aléatoires possiblement interdépendants comme contenus dans le modèle *GSPRE*. Ainsi on voit que l'hypothèse nulle est rejetée avec un niveau de confiance de 99%. Nous en concluons donc que notre modèle empirique doit contenir un terme traduisant soit l'interdépendance spatiale, soit des effets aléatoires spatialement interdépendants ou possiblement les deux. Afin de véritablement justifier la présence du terme $\rho W \ln(FDI)$, nous testons l'hypothèse nulle $\rho = 0$ en supposant qu'il est possible qu'il y aie des effets aléatoires spatialement interdépendants, $\sigma_\mu^2 \geq 0$. Encore une fois, nous rejetons l'hypothèse nulle comme le montre la seconde ligne du tableau 5.1 et ce à un niveau de confiance de 90%. Il est à noter que ce niveau est jugé plus qu'acceptable étant donné le faible nombre d'observations dont nous disposons. Pour écarter des effets aléatoires provinciaux interdépendants spatialement, nous utilisons le test d'Hausman adapté au contexte spatial proposé par Mutl et Pfaffermayr dans [61] et implanté dans le progiciel *splm* du logiciel *R*. La méthode d'estimation des moments généralisée est utilisée pour ce test, car celle-ci permet de ne pas tenir compte de la forme du terme d'erreur. Toujours en utilisant le modèle de base donné par l'équation 4.2, nous obtenons avec un degré de liberté de 4 une valeur de la statistique χ -carré de 10,2251 et une valeur de 0,0368 comme valeur critique associée (*P*-value). L'hypothèse nulle de ce test est que les effets aléatoires constituent la bonne spécification. Nous la rejetons donc avec un niveau de confiances de 95% et considérons des effets fixes provinciaux.

2. Voir [60] pour plus de renseignement sur ce progiciel.

Par la suite, nous estimons les modèles *SDM*, *SAC*, *SAR*, *GSPRE* et *SEM* donnés par les équations 5.1, 4.2, 2.12, 5.2 et 2.16 respectivement. Nous calculons également pour chacun de ceux-ci le critère d'information d'Akaike (CAI). Ce critère donne une mesure de qualité d'ajustement du modèle. Rappelons que R^2 n'est pas adéquat en général pour des modèles empiriques spatiaux et plus particulièrement pour notre modèle puisque celui-ci contient des termes d'erreur qui sont spatialement interdépendants. Une autre mesure comme la valeur du logarithme du maximum de vraisemblance peut être utilisée. Nous soulignons cependant que cette statistique augmente avec le nombre de variables incluses dans le modèle se rendant ainsi vulnérable au problème de surestimation. Cette problématique peut être contournée avec le critère d'information d'Akaike. Les résultats de nos estimations sont contenus dans le tableau B.1. La dernière ligne est celle dont nous nous préoccupons le plus puisqu'elle contient les valeurs du CIA. Une première constatation est que le CIA des colonnes 1 et 4, modèle *SDM* et *SAR* avec des effets fixes provinciaux, sont beaucoup plus faibles que ceux donnés dans les colonnes 2 et 4 respectivement confirmant ainsi le test spatial d'Hausman précédemment effectué. Afin d'écarter définitivement les effets aléatoires, le modèle *GSPRE* est estimé. Le CIA obtenu pour ce modèle est de 31,735 ce qui est beaucoup plus grand que celui du modèle *SDM* et *SAR* de 4,584 et 20,012 respectivement.

Finalement, on remarque que le modèle *SEM* contient la valeur la plus faible de CIA. Cependant, ce modèle ne fournit aucune estimation statistiquement significative. Ceci concorde avec les deux tests de multiplicateur de Lagrange qui confirmait la présence du terme $\rho W \ln(FDI)$. Avec ce modèle écarté, celui ayant la plus faible valeur du CIA est le modèle *SAC* avec une valeur de 4,547. Bien que légèrement inférieur à celle du modèle *SDM* avec des effets fixes provinciaux, le terme d'autocorrélation spatial $\rho W \ln(FDI)$ est statistiquement significatif seulement pour le modèle *SAC*. Nous choisissons donc le modèle *SAC* en nous basant sur les résultats des tests d'autocorrélation de multiplicateur de Lagrange.

5.2 Modèle d'investissement et variables de contrôle

Dans cette section, nous nous efforçons de déterminer quel est l'impact de l'ajout des variables de contrôle au modèle de base présenté par l'équation 4.2 et ne contenant que le coût unitaire et le *PIB* comme variables explicatives de la province hôte. Ces différentes variantes du modèle de base nous donne l'occasion de vérifier si le modèle d'investissement (Section 2.3.2) est robuste ou non au sujet de l'interdépendance spatiale.

Le tableau B.2 et B.3 présentent les résultats des estimations de diverses variantes d'inclusion de variables de contrôle sur l'ensemble de nos données de panel. Tous deux, construits de la même manière, nous présentent dans la première colonne l'estimation par les moindres carrés de l'équation 4.1. La colonne deux donne les résultats du modèle *SAC* appelé modèle de base et sert de point de comparaison pour les autres estimations spatiales. Les colonnes 3 et 4 introduisent respectivement les variables de contrôle *NB_etudiant* et *NB_diplome* alors que la colonne 5 celle contrôlant la population. La dernière colonne présente les estimés de la spécification complète, c'est-à-dire que toutes les variables de contrôle sont incluses. Chacune de ces régressions y compris celle estimée par la méthode des moindres carrés utilise des effets fixes provinciaux. Les régressions des colonnes 2 à 6 sont quant à elles estimées par la méthode du maximum de vraisemblance. Nous mentionnons que nous appliquons la transformation orthogonale proposée par Lee et Yu dans [51] étant donné le nombre restreint de périodes de temps à notre disposition.

Premièrement, mentionnons que nous présentons les résultats de la colonne 1 du tableau B.2 à des fins de comparaison avec la littérature empirique de la théorie de la gravité. Les estimés que l'on y voit ont tous le signe prédit par la littérature. Pourtant, seul le coût unitaire est statistiquement significatif, et ce à 10%. La situation est semblable pour le tableau B.3 où le taux marginal effectif de taxation des entreprises est inclus comme variable explicative. Fait inusité, le signe du coefficient estimé de cette variable est positif ce qui a priori semble aller à l'encontre des prédictions théoriques. De plus, ceci contredit les résultats empiriques de Bensassy-Quere et col. [9] présentés à la section 2.2. Nous discuterons de ce point plus

en détail subséquemment. Un fait intéressant des régressions spatiales des colonnes 2 à 6 du tableau B.2 est que l'estimé de ρ pour chacune de ces régressions est statistiquement significatif. Cette caractéristique démontre donc une très grande robustesse de l'interdépendance spatiale face à l'utilisation des variables de contrôle. Cette constatation demeure même en contrôlant pour l'hétérogénéité des provinces, résultat également obtenu dans [71] et [32].³ L'importance de la dimension spatiale est supportée également par la significativité statistique à un niveau de confiance d'au moins 90 % du coefficient estimé de $\ln(\text{sur_market_c})$. Son signe est toujours positif peu importe les variables de contrôle utilisées. Le potentiel des marchés voisins est certainement l'un des facteurs prédominant d'après les résultats obtenus avec une valeur estimée allant de 20,132 sous le modèle de base (colonne 2) à 24,142 pour le modèle complet de la colonne 6. Ainsi, l'augmentation de 1% de la moyenne pondérée par les distances séparant les capitales provinciales des *PIB* provinciaux canadiens autres que la province hôte provoque une augmentation de plus de 20% de l'investissement direct américain. Cette interprétation n'est que partiellement vraie puisque le coefficient étant positif suggère que la variable $\ln(\text{sur_market_c})$ capte également un effet d'agglomération. Nos résultats témoignent donc d'un modèle d'investissement vertical complexe pour les capitaux américains investis et ce de façon très robuste.

Un autre point intéressant, pour ne pas le qualifier d'énigme, est le signe négatif qu'ont tous les coefficients estimés du *PIB* lorsque l'interdépendance spatiale est prise en considération. Le fait que le *PIB* soit en plus non statiquement significatif ajoute encore plus au mystère. Pour continuer avec les anomalies, le coefficient estimé pour la population est quant à lui positif alors que la théorie le prédisait de signe opposé. Une explication possible est de revoir l'interaction entre la population et l'investissement. Il serait plausible de considérer la population comme une proxy mesurant le marché potentiel de la province hôte poussant ainsi à revoir son signe attendu comme étant positif. Nous écartons cette possibilité étant donné que dans des résultats non publiés des mêmes régressions utilisant des variables per capita n'ont présenté aucune différence. Une cause plus sensée semble être l'omission d'une

3. Bloginien et col. ont cependant obtenu des résultats contraires. Il faut mentionner que ces auteurs ont également montré que l'échantillon avait une grande importance en ce qui a trait à la significativité du terme $\rho W \ln(FDI)$.

ou plusieurs variables corrélées avec le *PIB* et la population. Une autre explication réside peut être dans le cas particulier que nous étudions ici, c'est-à-dire la distribution spatiale particulière du *PIB*. En effet, l'Ontario et le Québec constituent les deux plus gros marchés. Pourtant, les investissements en Alberta et Colombie-Britannique sont assez près, en valeur, de ceux faits en sol québécois et ontarien comme le montre la figure 5.1. Cette observation

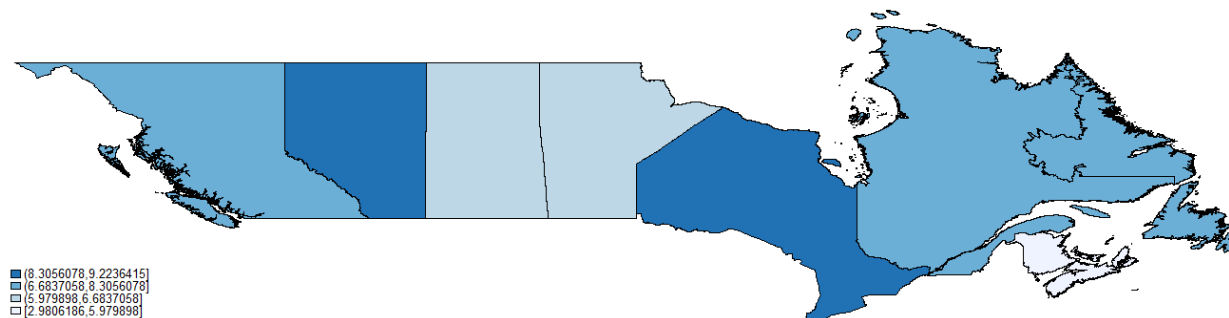


FIGURE 5.1 – Répartition de l'investissement direct américain dans les provinces canadiennes en 2011

coïncide avec le modèle d'investissement vertical complexe et l'effet d'agglomération observé précédemment. Ainsi, nos résultats nous permettent d'affirmer que la présence de ressources prime sur le marché potentiel de la province hôte lorsque la prise de décision concernant la localisation des investissements américains est abordée. Cette affirmation est lourde de sens puisque ceci laisse entrevoir que les investissements visent en premier lieu l'extraction et l'exploitation des ressources. Ainsi, certaines politiques gouvernementales pourraient être mises en place afin de favoriser le développement par exemple de la transformation afin d'inciter les multinationales à vendre leurs produits en sol canadien pour augmenter l'offre de produits et de créer une compétition plus grande sur le marché canadien. La question du marché visé par les investissements directs américains sera abordée dans la prochaine section.

Une analyse similaire peut être faite à partir du tableau B.3 lorsque l'on ajoute le taux marginal effectif de taxation des entreprises (METR). Fait marquant est que la valeur du coefficient de $\ln(\text{sur_market_c})$ a pratiquement doublé allant de 46,573 à 48,729. Cela confirme bien que cette variable capture l'effet d'agglomération et ne mesure pas uniquement le potentiel de marché, car une augmentation de 1% de la moyenne pondérée des *PIB* des provinces canadiennes ne peut produire à elle seule une augmentation de plus de 46% de l'in-

investissement direct étranger. Une valeur de 0,028 est estimée pour le coefficient de la variable *METR* pour le modèle de base (colonne 2). Toutes les autres spécifications obtiennent pour la même variable une valeur de 0,027. Ceci témoigne donc d'une très grande robustesse face aux variables de contrôle comme la population et la qualité de la main d'oeuvre. De plus, cette variable est statistiquement significative à un niveau de confiance de 99%. On remarque que le signe du coefficient estimé est positif pour toutes les spécifications utilisées. Comme mentionné à la section 3.2, cette variable est susceptible de capter également l'importance de l'accessibilité et de la qualité des infrastructures nécessaires à l'investissement. Le signe positif observé semble indiquer que c'est cet effet qui stimule l'investissement et domine celui des taxes imposées sur les multinationales. Il faut être prudent ici puisque nous utilisons des données agrégées ne révélant que l'effet global et dominant observé. Ainsi, des données désagrégées au niveau des industries seraient plus pertinentes compte tenu des différences entre les taux de taxations et les mesures fiscales pour les industries. Avec des données sur l'investissement sectoriel, il serait primordial d'utiliser des variables de contrôle pour l'accessibilité des ressources spécifiques nécessaires aux industries afin de contrôler pour l'effet d'agglomération qui peut également jouer un rôle sur la répartition spatiale de l'investissement. Les auteurs de [15] ont effectivement remarqué des changements dans leurs résultats lorsque des données désagrégées sont utilisées. Il va s'en dire que nous ne pouvons pas écarter la possibilité qu'une ou des variables omises soit la cause de cette corrélation positive entre la variable *METR* et l'investissement direct. Une candidate est la structure industrielle qui diffère entre les provinces. Le taux marginal effectif de taxation des entreprises varie énormément entre les différentes industries [28, TABLE 6A]. De ce fait, le taux moyen de taxation est nécessairement relié à la structure industrielle qui, elle, est corrélée l'investissement direct étranger. Finalement, l'abondance des ressources peut justifier la corrélation positive observée entre le *METR* et l'investissement direct étranger. Dans ce cas, les firmes ont une faible élasticité entre le stock de capital et le niveau de taxation. En d'autres mots, les gouvernements peuvent taxer d'avantage et ainsi la causalité entre le *METR* et l'investissement direct est en sens inverse.

Nous terminons l'analyse de ces deux tableaux en mentionnant que les valeurs de λ estimés sont toutes négatives et statistiquement significatives. Une motivation conjoncturelle de l'in-

sersion de $\lambda W\epsilon_t$ à la section 4.2 peut expliquer également ces résultats. Nous observons que l'investissement est négativement corrélé avec les chocs provinciaux temporels spatialement interdépendants. En particulier, l'implantation de programme spécifiquement conçus pour relancer une économie provinciale pousse vers le bas les investissements dans une province voisine. Cet effet diminue avec la distance séparant ces deux provinces. En supposant que les effets de ses politiques gouvernementales "spontanées" constituent la principale source de chocs interdépendants spatialement, il est certain que le manque de coordination de ces politiques cause des distorsions sur l'investissement direct américain. Avec une telle perspective, il serait prudent de mieux quantifier les effets de ces politiques sur les provinces afin de minimiser leurs impacts et de mieux coordonner leurs applications pour en tirer profit.

5.3 Marché visé

Le tableau B.4 et B.5 présentent les résultats des régressions faisant varier la mesure du potentiel des marchés voisins de la province hôte de l'investissement direct. La colonne 1 considère le marché canadien exclusivement et correspond au modèle de base déjà présenté précédemment. Les potentiels mesurés dans les colonnes 2 à 4 sont, dans l'ordre, ceux des marchés limitrophes américains exclusivement, marché canadien et limitrophes américains dans une même variable et pour terminer le marché canadien et limitrophes américains mesurés séparément. La dernière spécification est justifiée par l'utilisation du logarithme sur la variable de potentiel. Finalement, la dernière colonne donne les résultats d'une régression utilisant une différente mesure des distances entre les provinces afin de tester la robustesse de nos résultats sur ce point.

Un premier fait important qui ressort du tableau B.4 est que la proximité du marché américain est corrélée positivement avec l'investissement direct américain rejetant l'hypothèse que seul le marché canadien est visé par l'investissement direct des États-Unis. La régression utilisant uniquement le marché américain comme marché potentiel obtient des estimés de 4,233 pour la variable $\ln(\text{sur_only_usa})$. De plus, les régressions qui incluent le marché des états américains limitrophes de quelque façon que se soit donnent des valeurs

très proches les unes des autres sans être statistiquement identiques. Une augmentation de la moyenne pondérée du *PIB* des provinces canadiennes et des états américains considérés dans cette études provoque une hausse de 4,426 % de l'investissement direct américain. L'importance de la proximité du marché des États-Unis semble suggérer que les multinationales américaines choisissent d'investir au Canada en visant le pays d'origine. Ce genre de comportement concorde d'avantage avec le modèle d'investissement vertical pur dans lequel les firmes relocalisent certaines activités là où les coûts sont moins élevés. Ce type de modèle propose qu'une fois les activités terminées, le véritable marché est situé en dehors de ceux considérés par les régressions. Or dans le cas de cette étude, le marché visé est voisin des localisations des investissements. Ceci est confirmé à la colonne 4 par une valeur estimée statistiquement significative de 4,977 pour la variable $\ln(\text{sur_only_usa})$ alors que celle de la variable $\ln(\text{sur_market_c})$ n'est pas statistiquement significative. Cependant, nous réaffirmons que le modèle d'investissement observé est bel et bien de type vertical complexe comme le témoigne le signe positif de l'estimé de ρ et de toutes les variables mesurant le potentiel des marchés voisins. Nos observations proposent que des ressources essentielles aux investissements sont localisées dans les différentes provinces canadiennes et disponible à un coût moins élevé que celui au États-Unis. Ces ressources peuvent être naturelles ou encore être une main-d'oeuvre qualifiée et spécialisée reliée possiblement à certain secteur de l'économie. L'hypothèse de cette spécification sectorielle est appuyée par le signe positif des variables mesurant le potentiel des marchés voisins et témoignant ainsi d'un effet d'agglomération. Un deuxième point digne de mention est que l'interdépendance spatiale représentée par $\rho W \ln(FDI)$ demeure statistiquement significatif et ce à un niveau de confiance d'au moins 5%. Lorsque le marché américain est inclus, la valeur de ρ estimé varie très peu démontrant encore une fois la grande robustesse de nos résultats face à différentes mesures du potentiel des marchés voisins. Nous observons à nouveau que l'impact des chocs aléatoires provinciaux demeurent également statistiquement significatifs avec des valeurs légèrement plus élevées allant de -0,839 à -0,808.

Le tableau B.5 confirme les observations mentionnées au dernier paragraphe. On constate à la colonne 4 que les valeurs estimées pour les variables $\ln(\text{sur_market_c})$ et $\ln(\text{sur_only_usa})$

sont de 29,140 et 3,370 respectivement et statiquement significatives à 15%. En comparant cette colonne avec la colonne 4 du tableau B.4, on constate que la valeur de la variable $\ln(\text{sur_market_c})$ est maintenant positive et que le marché canadien pèse d'avantage dans la balance décisionnelle de la localisation de l'investissement. Un fait remarquable est que les valeurs estimées du coefficient du *METR* présentées dans le tableau B.5 varient entre 0,020 et 0,028. Ces valeurs sont très similaires à celles du tableau B.3 démontrant une robustesse de nos résultats face à diverses mesures du potentiel des marchés voisins.

Finalement, le changement dans les mesures des distances causé par l'utilisation des métropoles plutôt que ces capitales provoque des effets ambigus. Premièrement, tous les coefficients estimés dans la dernière colonne des tableaux B.4 et B.5 ont les signes prédits par la théorie. Pourtant le type de modèle d'investissement demeure tant qu'à lui inchangé. Cependant, une valeur d'à peine 0,003 et statistiquement significative est obtenue pour mesurer l'impact du potentiel de marché laissant ainsi présagé une quasi inexistence d'effet d'agglomération. De telles constatations semblent indiquer que nos résultats sont sensibles à la mesure de distance utilisée.

Chapitre 6

Conclusion

La littérature empirique portant sur l'investissement direct étranger et prenant en considération l'interdépendance spatiale a cru énormément durant la dernière décennie. Ce mémoire y contribue en étudiant l'investissement direct américain dans les provinces canadiennes entre 2004 et 2011. Nos résultats confirment la présence d'interdépendance spatiale tant par la significativité de la variable mesurant le potentiel des marchés voisins que par celles des variables spatialement retardées de l'investissement et des chocs aléatoires. De plus, le modèle d'investissement vertical complexe prédomine d'après nos résultats. Ce modèle est observé conjointement avec un effet d'agglomération important donnant ainsi lieu à des estimés reliés au potentiel des marchés voisins très élevés. Étonnamment, les régressions estimées démontrent une très grande robustesse quant au taux marginal effectif de taxation des entreprises. La surprise est d'autant plus grande que les estimations obtenues démontrent une corrélation positive entre le METR et l'investissement direct provenant des États-Unis. Ce résultat demeure mitigé puisque très peu d'étude utilise explicitement cette mesure dans une situation où les régions étudiées sont des provinces à l'intérieur d'un même pays. Finalement, nous obtenons que la proximité des états américains limitrophes au Canada est un facteur non négligeable. L'exclusion des marchés provinciaux canadiens dans le processus décisionnel de la localisation de l'investissement direct américain ne peut cependant pas être confirmée à partir de nos résultats. Seule une de nos régressions abonde dans ce sens.

Plus que jamais dans un contexte où la mondialisation est grandissante, les économies provin-

ciales canadiennes ne peuvent ignorer l'interdépendance spatiale à laquelle elles sont assujetties. Les politiques mises en place par les gouvernements provinciaux résonnent bien au-delà de leurs frontières respectives et s'entrechoquent créant ainsi des distorsions négatives en particulier sur l'investissement direct américain. Une meilleure coordination ne pourrait être que bénéfique. Le taux de taxation des entreprises est un exemple où la synchronisation des politiques provinciales serait profitable. Bien qu'au niveau international, une course vers le bas concernant la taxation des corporations a été entreprise depuis un certain temps, nos résultats laissent à penser qu'un tel comportement serait injustifié puisque le niveau de taxation des entreprises n'entraîne pas d'effets négatifs sur l'investissement direct étranger. L'accès aux ressources tant naturelles qu'humaines demeure un atout important qui se doit d'être maintenu à l'aide, entre autres choses, de taxes afin de stimuler l'investissement direct américain. D'autres défis attendent également les provinces d'après nos résultats. Une diversification des investissements avec une importance accrue accordée à ceux qui ont pour objectifs le marché canadien, et ce malgré la présence de notre voisin américain. De tels investissements ne font qu'augmenter la compétitivité au sein des marchés canadiens, mais apportent aussi de plus grandes retombées économiques et une croissance plus élevée.

Nous terminons en mentionnant que nos résultats ne témoignent que des courants dominants et des effets globaux étant donné le niveau d'agrégation de nos données. Il est indéniable que des données désagrégées au niveau des industries pour chacune des provinces permettraient d'atteindre un plus haut niveau de compréhension de l'interdépendance spatiale de l'investissement direct américain ainsi que des effets collatéraux engendrés par celle-ci au sein des économies provinciales.

Bibliographie

- [1] Andrew B. Abel and Janice C. Eberly, *A Unified Model of Investment under Uncertainty*, American Economic Review **84** (1994), no. 5, 1369–1384.
- [2] James E. Anderson, *The Gravity Model*, Annual Review of Economics **3** (2011), no. 1, 133–160.
- [3] James E. Anderson and Eric Van Wincoop, *Trade costs*, National Bureau of Economic Research, 2004.
- [4] Luc Anselin, *Spatial econometrics : methods and models*, Vol. 4, Kluwer Academic Pub, 1988.
- [5] Luc Anselin and Raymond Florax, *Small sample properties of tests for spatial dependence in regression models : Some further results*, New directions in spatial econometrics, 1995, pp. 21–74.
- [6] Badi H. Baltagi, Peter Egger, and Michael Pfaffermayr, *Estimating models of complex FDI : Are there third-country effects ?* Journal of Econometrics **140** (2007), no. 1, 260–281.
- [7] Badi H Baltagi, Peter Egger, and Michael Pfaffermayr, *A generalized spatial panel data model with random effects*, Working Paper, Syracuse University, 2007.
- [8] Badi H. Baltagi, Seuck Heun Song, and Won Koh, *Testing panel data regression models with spatial error correlation*, Journal of econometrics **117** (2003), no. 1, 123–150.
- [9] Agnes Benassy-Quere, Lionel Fontagne, and Amina Lahreche-Revil, *How Does FDI React to Corporate Taxation ?* International Tax and Public Finance **12** (2005), no. 5, 583–603.
- [10] Jeffrey H. Bergstrand, *The generalized gravity equation, monopolistic competition, and the factor-proportions theory in international trade*, The review of economics and statistics (1989), 143–153.
- [11] ———, *The Heckscher-Ohlin-Samuelson model, the Linder hypothesis and the determinants of bilateral intra-industry trade*, The Economic Journal **100** (1990), no. 403, 1216–1229.
- [12] Jeffrey H Bergstrand and Peter Egger, *A theoretical and empirical model of international trade and foreign direct investment with outsourcing : Part I, developed countries*, unpublished paper (2004).
- [13] Roger S. Bivand, Edzer J. Pebesma, and Virgilio Gómez Rubio, *Applied spatial data : analysis with R*, Springer, 2008.
- [14] Bruce A. Blonigen and Ronald B. Davies, *The effects of bilateral tax treaties on US FDI activity*, International Tax and Public Finance **11** (2004), no. 5, 601–622.

- [15] Bruce A. Blonigen, Ronald B. Davies, Glen R. Waddell, and Helen T. Naughton, *FDI in space : Spatial autoregressive relationships in foreign direct investment*, *European Economic Review* **51** (2007), no. 5, 1303–1325.
- [16] Bruce A. Blonigen and Miao Wang, *Inappropriate pooling of wealthy and poor countries in empirical FDI studies*, National Bureau of Economic Research, 2004.
- [17] Richard Blundell, Stephen Bond, Michael Devereux, and Fabio Schiantarelli, *Investment and Tobin's Q : Evidence from company panel data*, *Journal of Econometrics* **51** (1992), no. 1, 233–257.
- [18] Tullio Bucciato and Francesco Santangelo, *Foreign direct investments distribution in the Russian Federation : do spatial effects matter ?* (2009).
- [19] Statistics Canada, *Foreign and Domestic Investment in Canada*, (Various) Tables **61-232** (2004/06).
- [20] ———, *Foreign and Domestic Investment in Canada*, (Various) Tables **61-232** (2006/08).
- [21] ———, *Foreign and Domestic Investment in Canada*, (Various) Tables **61-232** (2007/09).
- [22] ———, *Foreign and Domestic Investment in Canada*, (Various) Tables **61-232** (2008/10).
- [23] ———, *Foreign and Domestic Investment in Canada*, (Various) Tables **61-232** (2009/11).
- [24] Duanjie Chen, Yvan Guillemette, Jack M. Mintz, and Finn Poschmann, *Tax Competitiveness Report* (2005).
- [25] Duanjie Chen and Jack M. Mintz, *Taxing business investments : A new ranking of effective tax rates on capital*, World Bank (2008).
- [26] ———, *Canada's Tax Competitiveness After a Decade of Refoms : Still an Unfinished Plan*, School of Public Policy, SPP Research Papers, University of Calgary **3** (2010), no. 5.
- [27] ———, *Federal-Provincial Business Tax Reforms : A Growth Agenda with Competitive Rates and a Neutral Treatment of Business Activities*, School of Public Policy, SPP Research Papers, University of Calgary **4** (2011), no. 1.
- [28] ———, *2012 Annual Global Tax Competitiveness Ranking- Canadian Good News Story*, School of Public Policy, SPP Research Papers, University of Calgary **5** (2012), no. 28.
- [29] Cletus C. Coughlin and Eran Segev, *Foreign direct investment in China : a spatial econometric study*, *The World Economy* **23** (2000), no. 1, 1–23.
- [30] Serge Coulombe, *The Canada-U.S. Productivity Growth Paradox*, Productivity issues in canada, 2002.
- [31] D. Drukker, Ingmar R. Prucha, and R. Raciborski, *Maximum-Likelihood and Generalized Spatial Two-Stage Least-Squares Estimators for a Spatial-Autoregressive Model with Spatial-Autoregressive Disturbances*, *Stata Journal* (forthcoming).
- [32] Peter Egger, Michael Pfaffermayr, and Hannes Winner, *An unbalanced spatial panel data approach to US state tax competition*, *Economics Letters* **88** (2005), no. 3, 329–335.

- [33] Karolina Ekholm, Rikard Forslid, and James R Markusen, *Export-Platform Foreign Direct Investment*, Journal of the European Economic Association **5** (2007), no. 4, 776–795.
- [34] Bulent Esiyok and Mehmet Ugur, *Foreign direct investment in provinces : A spatial regression approach to FDI in Vietnam* (2011).
- [35] Harry Garretsen and Jolanda Peeters, *FDI and the relevance of spatial linkages : do third-country effects matter for Dutch FDI?* Review of World Economics **145** (2009), no. 2, 319–338.
- [36] John P. Gould, *Adjustment Costs in the Theory of Investment of the Firm*, The Review of Economic Studies **35** (1968), no. 1, 47–55.
- [37] Zahoor Haq, Karl D. Meilke, and John Cranfield, *The Gravity Model and the Problem of Zeros in Agrifood Trade*, Canadian Agricultural Trade Policy Research Network Working Papers (2011).
- [38] Fumio Hayashi, *Tobin's marginal q and average q : A neoclassical interpretation*, Econometrica : Journal of the Econometric Society (1982), 213–224.
- [39] Elhanan Helpman, *A simple theory of international trade with multinational corporations*, The Journal of Political Economy (1984), 451–471.
- [40] Elhanan Helpman, Marc J. Melitz, and Stephen R. Yeaple, *Export versus FDI*, National Bureau of Economic Research, 2003.
- [41] Eunsuk Hong, *Location of foreign direct investment in China : A spatial dynamic panel data analysis by country of origin*, Ph.D. Thesis, 2008.
- [42] Dale W. Jorgenson, *Capital theory and investment behavior*, The American Economic Review **53** (1963), no. 2, 247–259.
- [43] Mudit Kapoor, Harry H. Kelejian, and Ingmar R. Prucha, *Panel data models with spatially correlated error components*, Journal of Econometrics **140** (2007), no. 1, 97–130.
- [44] Harry H Kelejian and Ingmar R Prucha, *A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances*, The Journal of Real Estate Finance and Economics **17** (1998), no. 1, 99–121.
- [45] Jörn Kleinert and Farid Toubal, *Gravity for FDI*, Review of International Economics **18** (2010), no. 1, 1–13.
- [46] Svetlana Ledyeva, *Spatial Econometric Analysis of Foreign Direct Investment Determinants in Russian Regions*, World Economy **32** (2009), no. 4, 643–666.
- [47] Svetlana Ledyeva and Mikael Linden, *Testing for foreign direct investment. gravity model for russian regions*, Electronic Publications of the University of Joensuu. [http ://joypub. joensuu. fi/publications/other_publications/ledyaeva_investment/ledyaeva. pdf](http://joypub.joensuu.fi/publications/other_publications/ledyaeva_investment/ledyaeva.pdf) (2006).
- [48] Lung-fei Lee, *Generalized method of moments estimation of spatial autoregressive processes*, Manuscript, Department of Economics, OSU (2001).

- [49] ———, *GMM and 2SLS estimation of mixed regressive, spatial autoregressive models*, *Journal of Econometrics* **137** (2007), no. 2, 489–514.
- [50] Lung-fei Lee and Xiaodong Liu, *Efficient GMM estimation of high order spatial autoregressive models with autoregressive disturbances*, *Econometric Theory* **26** (2010), no. 01, 187–230.
- [51] Lung-fei Lee and Jihai Yu, *Estimation of spatial autoregressive panel data models with fixed effects*, *Journal of Econometrics* **154** (2010), no. 2, 165–185.
- [52] ———, *Estimation of spatial panels*, Now Publishers Inc, 2011.
- [53] James P. LeSage, *Spatial econometrics*, Regional Research Institute, West Virginia University, 1999.
- [54] James P. LeSage and R. Kelley Pace, *Models for spatially dependent missing data*, *The Journal of Real Estate Finance and Economics* **29** (2004), no. 2, 233–254.
- [55] ———, *Spatial and spatiotemporal econometrics*, Elsevier JAI, 2004.
- [56] ———, *Introduction to spatial econometrics*. CRC Press, Taylor and Francis Group, 2009.
- [57] Robert E. Lucas, *Adjustment costs and the theory of supply*, *The Journal of Political Economy* **75** (1967), no. 4, 321–334.
- [58] James R. Markusen, *Multinationals, multi-plant economies, and the gains from trade*, *Journal of international economics* **16** (1984), no. 3, 205–226.
- [59] Marc J. Melitz, *The impact of trade on intra-industry reallocations and aggregate industry productivity*, *Econometrica* **71** (2003), no. 6, 1695–1725.
- [60] Giovanni Millo and Gianfranco Piras, *splm : Spatial Panel Data Models in R*, *Journal of Statistical Software* **47** (2012), no. 1, 1–38.
- [61] Jan Mutl and Michael Pfaffermayr, *The Hausman test in a Cliff and Ord panel model*, *The Econometrics Journal* **14** (2011), no. 1, 48–76.
- [62] Anh Nguyen and Thang Nguyen, *Foreign direct investment in Vietnam : An overview and analysis the determinants of spatial distribution across provinces*, Available at SSRN 999550 (2007).
- [63] Uwaoma G. Nwaogu, *Essays on Spatial Analysis of Foreign Direct Investment and Economic Growth Determinants in Developing Countries*, Ph.D. Thesis, 2012.
- [64] Lev Semenovich Pontriagin, *The mathematical theory of optimal processes*, Vol. 4, CRC PressI Llc, 1962.
- [65] Shalendra Sharma, Miao Grace Wang, and MC Wong, *A Spatial Analysis of Aggregate and Industry-Level FDI in China*, submitted paper (2010).
- [66] Lawrence H. Summers, Barry P. Bosworth, James Tobin, and Philip M. White, *Taxation and corporate investment : A q-theory approach*, *Brookings Papers on Economic Activity* **1981** (1981), no. 1, 67–140.
- [67] James Tobin, *A general equilibrium approach to monetary theory*, *Journal of money, credit and banking* **1** (1969), no. 1, 15–29.

- [68] Arthur B. Treadway, *On rational entrepreneurial behaviour and the demand for investment*, The Review of Economic Studies **36** (1969), no. 2, 227–239.
- [69] Hirofumi Uzawa, *Time preference and the Penrose effect in a two-class model of economic growth*, The Journal of Political Economy **77** (1969), no. 4, 628–652.
- [70] Swarnim Wagle, *Investing across borders with heterogeneous firms : do FDI-specific regulations matter ?* World Bank Policy Research Working Paper Series, Vol (2011).
- [71] Wei Wang, *Three Essays on Spatial Econometric Models with Missing Data*, Ph.D. Thesis, 2010.
- [72] Wei Wang and Lung-fei Lee, *Estimation of spatial panel data models with randomly missing data in the dependent variable*, Regional Science and Urban Economics **43** (2013), no. 3, 521 –538.
- [73] Stephen Ross Yeaple, *The complex integration strategies of multinationals and cross country dependencies in the structure of foreign direct investment*, Journal of International Economics **60** (2003), no. 2, 293–314.

Annexe A

Liste des états américains limitrophes du Canada

- Idaho
- Maine
- Michigan
- Minnesota
- Montana
- New Hampshire
- Newyork
- North Dakota
- Ohio
- Pensnsylvania
- Vermont
- Washington

Annexe B

Tableaux de résultats des régressions

Variables	SDM (fixed)	SDM (random)	SAC	SAR (fixed)	SAR (random)	GSPRE	SEM
Coût unitaire	0,505 (1,033)	0,263 (1,004)	-1,742** (0,766)	-1,715 (1,061)	-1,433** (0,709)	-0,681 (1,147)	-1,296 (1,147)
ln(PIB)	0,009 (1,477)	1,022*** (0,131)	-1,434 (1,269)	-0,971 (1,453)	0,971*** (0,111)	0,896*** (0,089)	-0,741 (0,089)
ln(Marché voisin) (canadien)	79,335** (34,838)	-4,820* (2,755)	20,132* (10,139)	13,393 (13,796)	-4,638* (2,461)	-6,025*** (1,785)	5,221 (1,785)
Constante		-44,098* (23,564)			4,305 (5,351)	9,212** (3,73)	
W(Coût unitaire)	-6,937*** (1,995)	-6,444*** (1,923)					
W(ln(PIB))	-12,97** (5,65)	1,453† (0,962)					
W(ln(sur_market_c))	47,037 (37,108)	19,925** (8,563)					
W(ln(FDI))	-0,320 (0,23)	-0,077 (0,212)	0,611*** (0,114)	0,319** (0,157)	0,267* (0,148)		
λ			-0,940*** (0,289)			0,408** (0,182)	0,328 (0,182)
ϕ						-1,225† (0,761)	
Critère d'information d'Akaike	4,584	42,756	4,547	20,012	50,748	31,735	-6,601

Significance levels : † : 15% * : 10% ** : 5% *** : 1%, Écart type échantillonale entre parenthèse

Tableau B.1 – Analyse spatiale de l'investissement direct américain - Sélection du modèle

Variables	OLS	Modèle de base		Étudiant		Population	Modèle complet
		Étudiants actuels	Diplômés	Étudiants actuels	Diplômés		
Coût unitaire	-1,965* (1,006)	-1,742** (0,766)	-1,954** (0,196)	-1,995* (1,008)	-2,068** (0,902)	-1,955** (0,900)	
$\ln(PIB)$	0,310 (1,084)	-1,434 (1,269)	-2,041† (0,813)	-1,628 (1,055)	-1,739 (1,390)	-2,041 (1,399)	
$\ln(sur_market_c)$		20,132* (10,139)	24,140** (1,395)	20,845* (1,442)	21,929** (10,805)	24,142** (10,897)	
$\ln(NB_etudiant)$	0,132 (0,376)		0,264 (10,888)			0,264 (0,254)	
$\ln(NB_diplome)$				0,379 (10,856)			
$\ln(Population)$	-1,129 (2,034)				1,003 (1,242)	0,007 (1,555)	
Constante	6,378 (11,863)						
$W(\ln(FDI))$		0,611*** (0,114)	0,634*** (0,115)	0,628*** (0,124)	0,629*** (0,116)	0,634*** (0,115)	
λ		-0,940*** (0,289)	-1,045*** (0,298)	-0,988*** (0,324)	-1,039*** (0,313)	-1,045*** (0,306)	

Significance levels : † : 15% * : 10% ** : 5% *** : 1% Écart type échantillonale entre parenthèse

Tableau B.2 – Analyse spatiale de l'investissement direct américain - Variables de contrôles excluant METR

Variables	OLS	Modèle de base		Étudiant		Population	Modèle complet
		Étudiants actuels	Diplômés	Étudiants actuels	Diplômés		
Coût unitaire	-0,146 (1,213)	-0,934 (0,875)	-1,243 (0,958)	-1,562 (1,125)	-1,221 (1,055)	-1,200 (1,042)	
$\ln(PIB)$	1,261 (1,266)	-1,864 [†] (1,21)	-2,371* (1,358)	-2,283 [†] (1,361)	-2,008 [†] (1,324)	-2,373* (1,358)	
<i>ME</i> <i>TR</i>	0,025*** (0,008)	0,028*** (0,007)	0,027*** (0,008)	0,027*** (0,008)	0,027*** (0,008)	0,027*** (0,008)	
$\ln(sur_market_c)$		46,952*** (12,56)	48,587*** (13,34)	47,253*** (13,012)	46,573*** (13,305)	48,729*** (13,434)	
$\ln(NB_etudiant)$	0,065 (0,366)		0,257 (0,214)			0,274 (0,27)	
$\ln(NB_diplome)$				1,042 (1,054)			
$\ln(Population)$	-0,281 (2,285)				0,871 (1,479)	-0,188 (1,775)	
Constante	-6,843 (14,466)						
$W(\ln(FDI))$		0,211 (0,188)	0,232 (0,199)	0,272 (0,200)	0,233 (0,202)	0,230 (0,201)	
λ		-0,639** (0,305)	-0,740** (0,336)	-0,783** (0,347)	-0,719** (0,352)	-0,732** (0,344)	

Significance levels : † : 15% * : 10% ** : 5% *** : 1% Écart type échantillonale entre parenthèse

Tableau B.3 – Analyse spatiale de l'investissement direct américain - Variables de contrôles avec ME*TR*

Variables	Canada	USA	Canada et USA		Métropole
			combinés	séparé	
Coût unitaire	-1,742** (0,766)	-1,010 (0,792)	-1,268† (0,808)	-0,715 (0,919)	-1,172† (0,744)
$\ln(PIB)$	-1,434 (1,269)	-1,831* (1,037)	-2,053* (1,131)	-1,314 (1,329)	1,005 (1,073)
$\ln(sur_market_c)$	20,132* (10,139)			-9,130 (14,686)	
$\ln(sur_can_usa)$			4,426*** (1,542)		
$\ln(sur_only_usa)$		4,233*** (1,412)		4,977** (1,902)	
$\ln(sur_market_metro)$					0,749*** (0,002)
$W(\ln(FDI))$	0,611*** (0,114)	0,407** (0,178)	0,438** (0,170)	0,391** (0,185)	0,664*** (0,118)
λ	-0,940*** (0,289)	-0,839*** (0,315)	-0,870*** (0,309)	-0,808** (0,327)	-0,822*** (0,296)

Significance levels : † : 15% * : 10% ** : 5% *** : 1% Écart type échantillonale entre parenthèse

Tableau B.4 – Analyse spatiale de l'investissement direct américain - Marché visé excluant METR

Variables	Canada	USA	Canada et USA		Métropole
			combinés	séparé	
Coût unitaire	-0,934 (0,875)	0,092 (0,905)	-0,253 (0,925)	-0,620 (0,978)	-0,055 (0,921)
$\ln(PIB)$	-1,864 [†] (1,21)	-0,953 (1,221)	-1,349 (1,285)	-1,947 [†] (1,296)	0,546 (1,225)
<i>METR</i>	0,028*** (0,007)	0,020*** (0,007)	0,021*** (0,007)	0,026*** (0,008)	0,023*** (0,008)
$\ln(sur_market_c)$	46,952*** (12,56)			29,140 [†] (17,465)	
$\ln(sur_can_usa)$			6,346*** (2,097)		
$\ln(sur_only_usa)$		5,771*** (1,949)		3,370 [†] (2,236)	
$\ln(sur_market_metro)$					0,003 [†] (0,002)
$W(\ln(FDI))$	0,211 (0,188)	0,043 (0,272)	0,061 (0,261)	0,034 (0,252)	0,400** (0,193)
λ	-0,639** (0,305)	-0,431 (0,376)	-0,472 (0,365)	-0,497 (0,351)	-0,659** (0,322)

Significance levels : † : 15% * : 10% ** : 5% *** : 1% Écart type échantillonale entre parenthèse

Tableau B.5 – Analyse spatiale de l'investissement direct américain - Marché visé avec METR

Annexe C

Tableaux descriptifs supplémentaires

Variables	Moyenne	Écart type	Min.	Max.	Obs.
Investissement	9 893,813	1 435,085	7 205,7	11 745,9	8
METR	22,171	4,941	17,7	31,4	7
Nombre d'étudiants	0,161	0,017	0,14	0,189	8
Nombre de diplômés	1 597,449	139,196	1 377,783	1 775,758	8
Coût unitaire	0,682	0,086	0,542	0,775	8
Population	3,533	0,194	3,239	3,778	8
PIB	174 040,6	9 411,203	157 088,1	187 037,1	8
Marché can.	527,238	19,577	493,526	555,239	8
Marché amé. et can. (combinés)	1 882,628	49,246	1 795,293	1 942,677	8
Marché amé.	1 355,39	31,695	1 301,767	1 389,077	8

Tableau C.1 – Statistiques descriptives pour la province de l'Alberta

Variabes	Moyenne	Écart type	Min.	Max.	Obs.
Investissement	3 763,075	417,548	2 955,6	4 353,3	8
METR	28,886	7,626	18,5	39,1	7
Nombre d'étudiants	0,227	0,024	0,203	0,267	8
Nombre de diplômés	1 811,128	116,223	1 645,208	1 974,233	8
Coût unitaire	0,685	0,037	0,627	0,725	8
Population	4,357	0,156	4,155	4,577	8
PIB	148 075,1	7 100,331	135 020,9	157 525	8
Marché can.	462,304	18,748	429,475	489,079	8
Marché amé. et can. (combinés)	3 221,461	129,042	2 978,864	3 358,653	8
Marché amé.	2 759,157	110,73	2 549,388	2 869,574	8

Tableau C.2 – Statistiques descriptives pour la province de la Colombie Britannique

Variabes	Moyenne	Écart type	Min.	Max.	Obs.
Investissement	693,55	159,314	450,5	886,6	8
METR	33,857	5,099	27	40,6	7
Nombre d'étudiants	0,051	0,005	0,047	0,062	8
Nombre de diplômés	456,41	27,231	428,292	498,742	8
Coût unitaire	0,665	0,038	0,62	0,723	8
Population	1,205	0,028	1,174	1,252	8
PIB	38 173,55	2 190,895	34 700,8	40 699,5	8
Marché can.	762,91	27,366	715,618	802,825	8
Marché amé. et can. (combinés)	2 793,449	62,054	2 696,307	2 872,889	8
Marché amé.	2 030,539	39,688	1 963,114	2 072,035	8

Tableau C.3 – Statistiques descriptives pour la province du Manitoba

Variabes	Moyenne	Écart type	Min.	Max.	Obs.
Investissement	336,488	82,344	257,1	501,6	8
METR	14,157	9,179	2,9	26,5	7
Nombre d'étudiants	0,032	0,001	0,031	0,033	8
Nombre de diplômés	295,031	7,011	286,475	302,808	8
Coût unitaire	0,705	0,052	0,641	0,78	8
Population	0,749	0,003	0,745	0,755	8
PIB	21 229,58	645,096	20 185,9	22 011,4	8
Marché can.	1 364,097	42,939	1 292,98	1 426,074	8
Marché amé. et can. (combinés)	4 339,795	97,398	4 172,797	4 457,424	8
Marché amé.	2 975,698	63,341	2 879,817	3 047,157	8

Tableau C.4 – Statistiques descriptives pour la province du Nouveau-Brunswick

Variabes	Moyenne	Écart type	Min.	Max.	Obs.
Investissement	690,5	299,831	356,9	1 116,5	8
METR	19,457	7,99	8	28,6	7
Nombre d'étudiants	0,052	0,001	0,05	0,055	8
Nombre de diplômés	369,891	12,123	358,55	387,333	8
Coût unitaire	0,704	0,044	0,642	0,77	8
Population	0,94	0,004	0,936	0,948	8
PIB	26 412,46	772,069	25 250,2	27 349,1	8
Marché can.	964,764	30,53	913,22	1 008,195	8
Marché amé. et can. (combinés)	3 400,455	76,187	3 272	3 488,941	8
Marché amé.	2 435,691	51,681	2 358,78	2 494,459	8

Tableau C.5 – Statistiques descriptives pour la province de la Nouvelle-Écosse

Variabes	Moyenne	Écart type	Min.	Max.	Obs.
Investissement	10 737,15	1 645,498	8 529,3	12 506,3	8
METR	31,9	9,512	18,7	43,4	7
Nombre d'étudiants	0,675	0,056	0,606	0,761	8
Nombre de diplômés	5 603,464	308,028	5 106,542	6 016,175	8
Coût unitaire	0,688	0,038	0,634	0,734	8
Population	12,871	0,34	12,391	13,366	8
PIB	480 518,7	12 491,63	458 650,2	496 953,9	8
Marché can.	536,901	20,495	503,787	566,371	8
Marché amé. et can. (combinés)	5 987,054	118,568	5 780,165	6 130,051	8
Marché amé.	5 450,154	111,423	5 240,643	5 590,203	8

Tableau C.6 – Statistiques descriptives pour la province de l'Ontario

Variabes	Moyenne	Écart type	Min.	Max.	Obs.
Investissement	3 929,138	961,673	2 441,9	4 879,6	8
METR	24,329	7,814	16,9	36,2	7
Nombre d'étudiants	0,467	0,019	0,451	0,502	8
Nombre de diplômés	3 202,095	142,022	2 967,217	3 401,258	8
Coût unitaire	0,672	0,033	0,626	0,718	8
Population	7,737	0,157	7,536	7,978	8
PIB	244 740,1	8 309,153	232 259,6	256 961,1	8
Marché can.	901,593	26,703	854,389	937,672	8
Marché amé. et can. (combinés)	4 699,589	105,634	4 525,633	4 809,085	8
Marché amé.	3 797,996	82,537	3 671,244	3 889,987	8

Tableau C.7 – Statistiques descriptives pour la province du Québec

Variabes	Moyenne	Écart type	Min.	Max.	Obs.
Investissement	789,1	199,731	517,2	1 146,9	8
METR	30,914	7,389	25	43,9	7
Nombre d'étudiants	0,039	0,01	0,024	0,053	8
Nombre de diplômés	394,869	23,218	368,633	424,283	8
Coût unitaire	0,582	0,073	0,481	0,67	8
Population	1,016	0,025	0,992	1,058	8
PIB	38 003,45	1 795,992	35 833,1	41 186,2	8
Marché can.	783,776	32,079	726,621	827,815	8
Marché amé. et can. (combinés)	2 422,459	63,537	2 312,646	2 504,054	8
Marché amé.	1 638,683	34,81	1 586,025	1 676,239	8

Tableau C.8 – Statistiques descriptives pour la province de la Saskatchewan

Variabes	Moyenne	Écart type	Min.	Max.	Obs.
Investissement	600,263	231,723	372,7	1 127,5	8
METR	15,657	3,99	11	21,1	7
Nombre d'étudiants	0,027	0,001	0,025	0,028	8
Nombre de diplômés	191,328	6,315	182,7	202,4	8
Coût unitaire	0,578	0,064	0,502	0,656	8
Population	0,511	0,004	0,506	0,517	8
PIB	17 064,36	948,859	15 763,2	18 221,1	8
Marché can.	544,835	17,445	515,026	569,763	8
Marché amé. et can. (combinés)	1 896,14	41,522	1 827,395	1 944,259	8
Marché amé.	1 351,306	27,996	1 308,07	1 384,153	8

Tableau C.9 – Statistiques descriptives pour la province de Terre-Neuve-Labrador

Annexe D

Figures supplémentaire

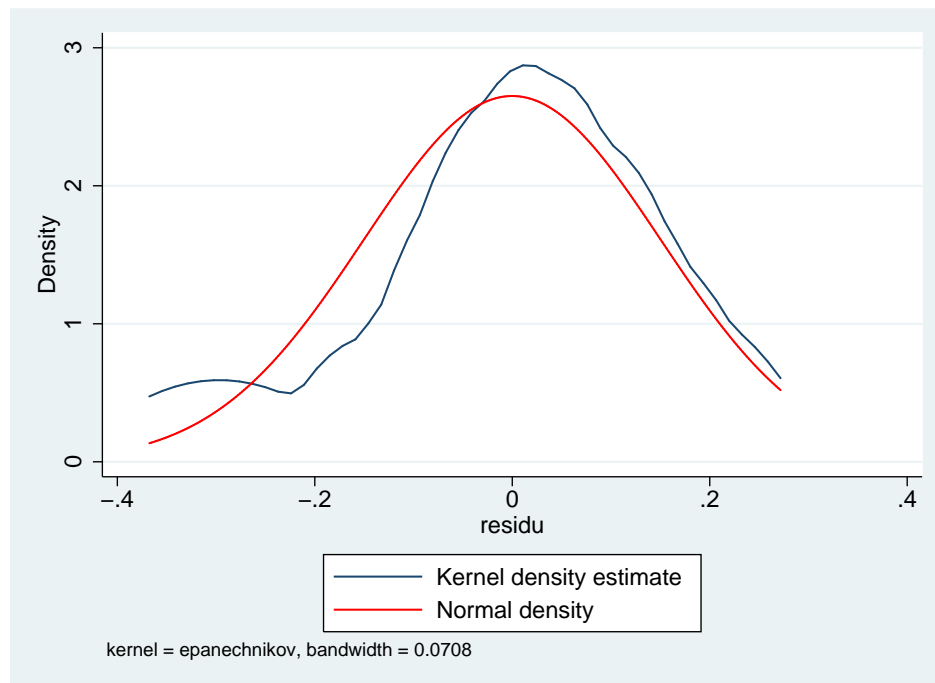


FIGURE D.1 – Distribution des résidus du modèle de base pour la province de l'Alberta

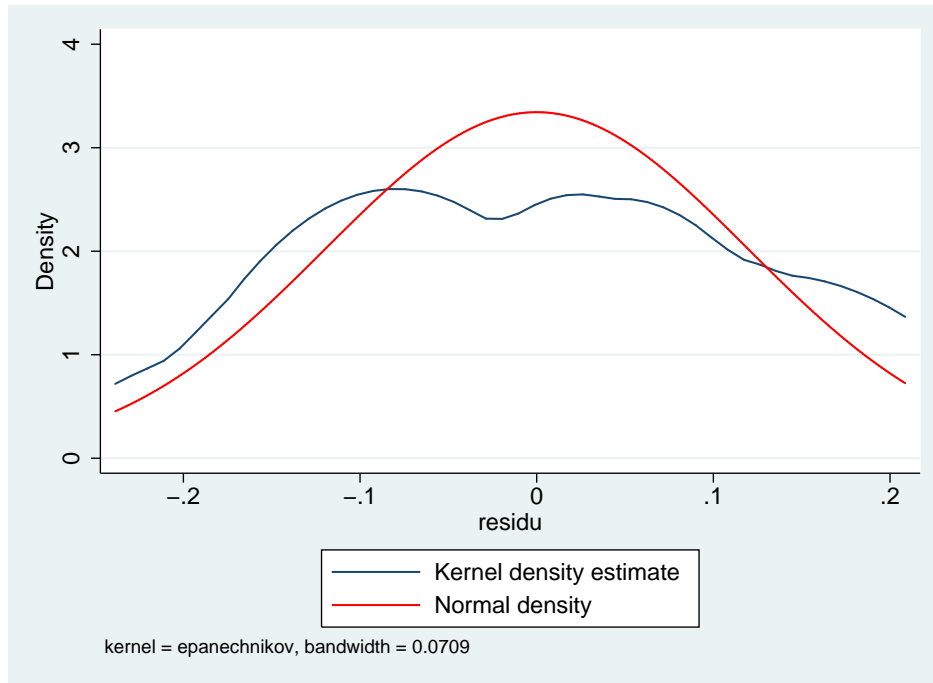


FIGURE D.2 – Distribution des résidus du modèle de base pour la province de la Colombie Britannique

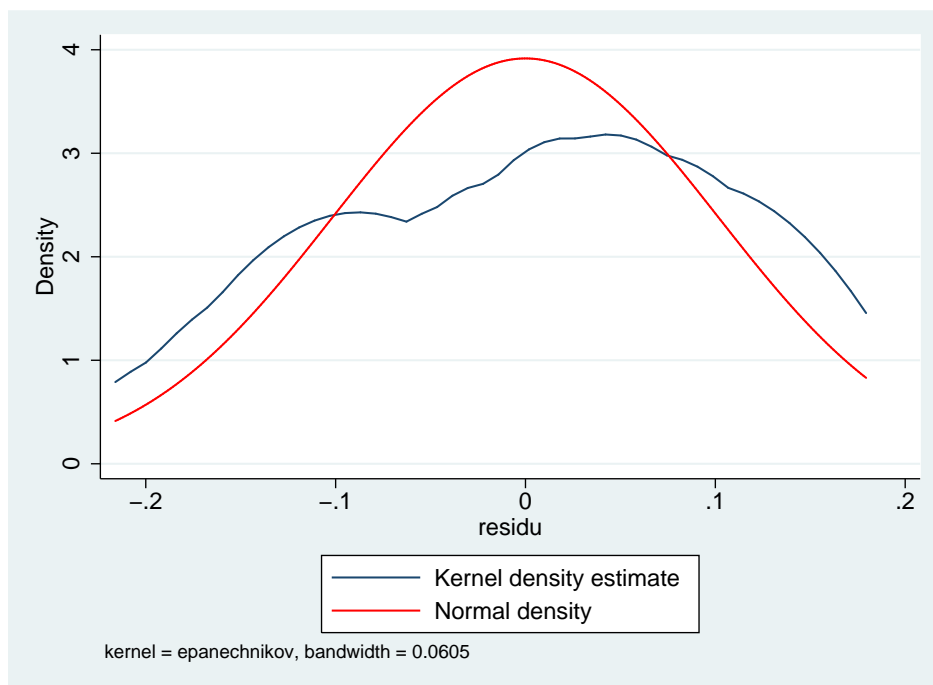


FIGURE D.3 – Distribution des résidus du modèle de base pour la province du Manitoba

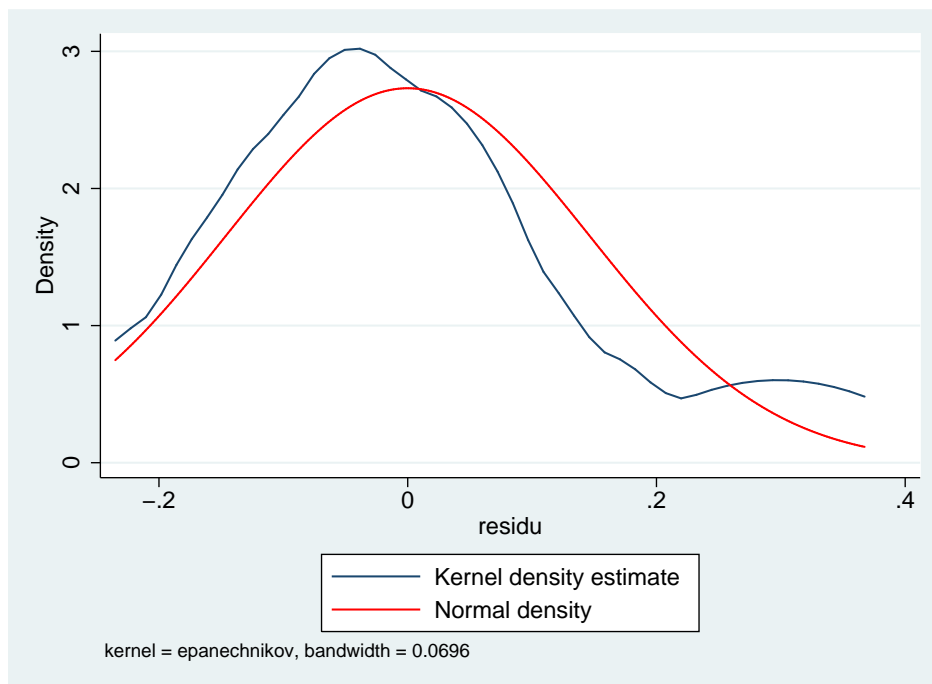


FIGURE D.4 – Distribution des résidus du modèle de base pour la province du Nouveau-Brunswick

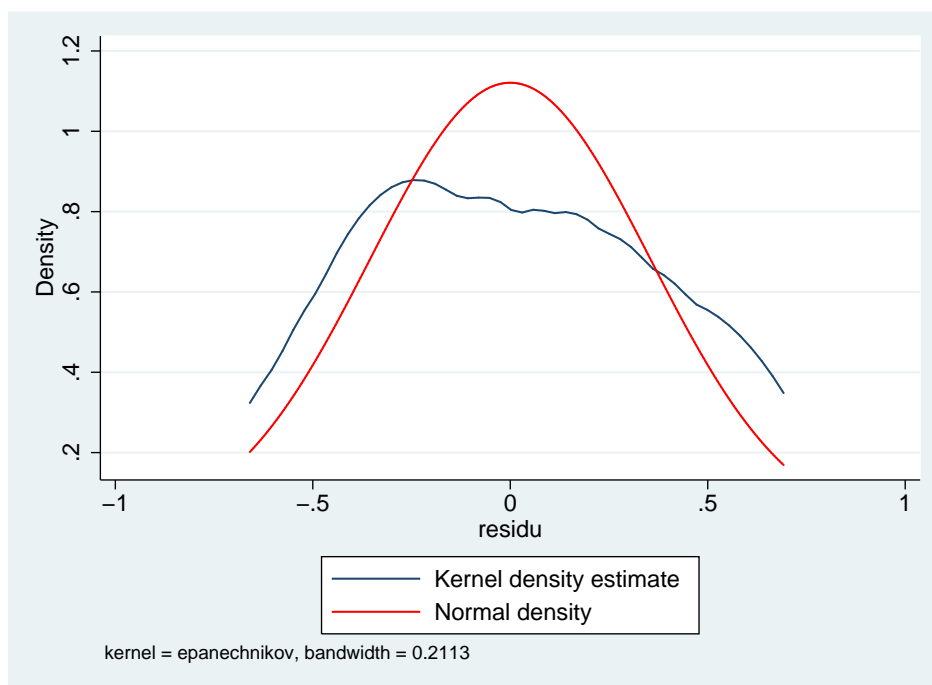


FIGURE D.5 – Distribution des résidus du modèle de base pour la province de la Nouvelle-Écosse

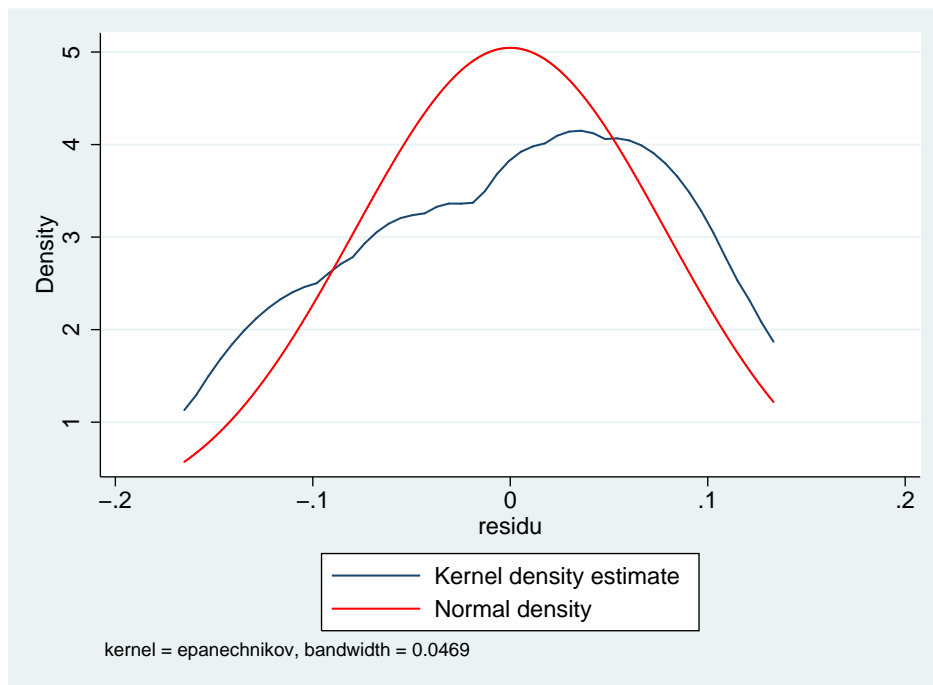


FIGURE D.6 – Distribution des résidus du modèle de base pour la province de l’Ontario

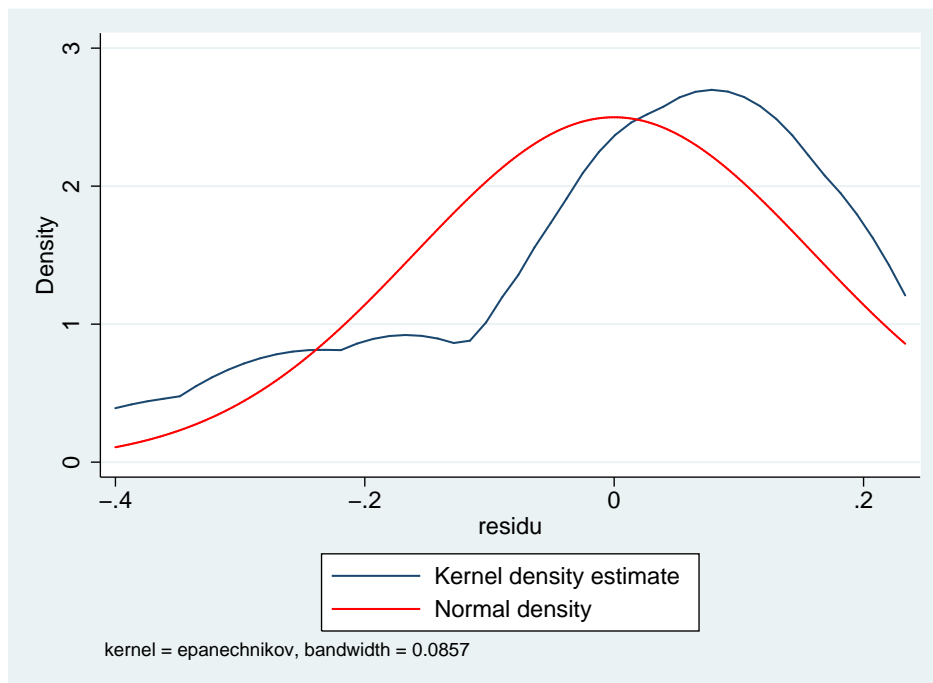


FIGURE D.7 – Distribution des résidus du modèle de base pour la province du Québec

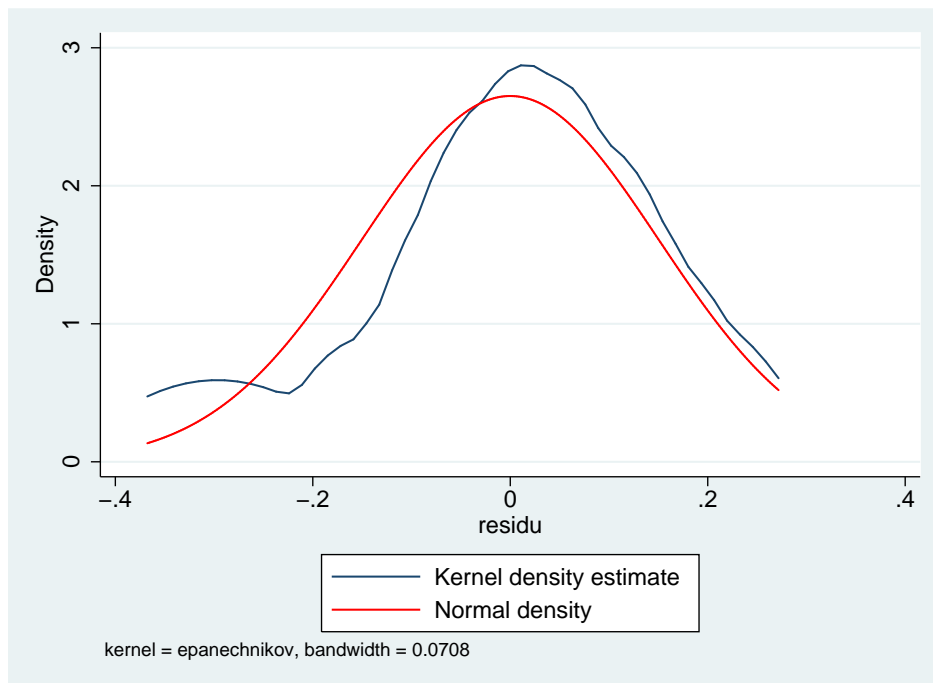


FIGURE D.8 – Distribution des résidus du modèle de base pour la province de la Saskatchewan

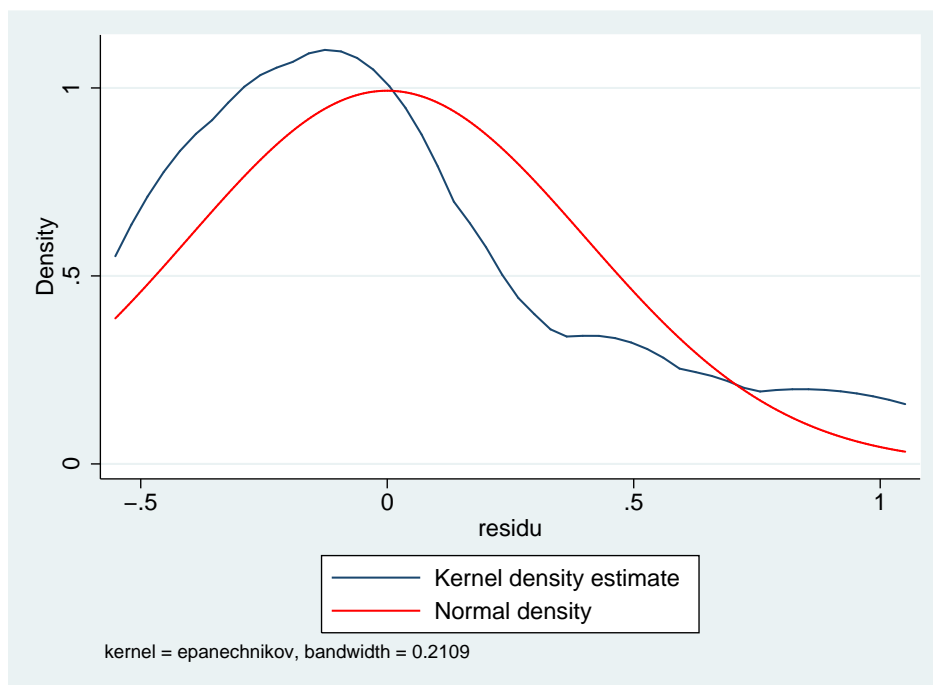


FIGURE D.9 – Distribution des résidus du modèle de base pour la province de Terre-Neuve-Labrador