

UNIVERSITE DE BOURGOGNE

Laboratoire d'Economie de Dijon (LEDi) UMR 6307 CNRS - U 1200 Inserm

Ecole doctorale LISIT 491 (Langages, Idées, Sociétés, Institutions, Territoires)

THÈSE

Pour obtenir le grade de

Docteur de l'Université de Bourgogne

Discipline : SCIENCES ECONOMIQUES

Par

SILEYMANE BA

Le 8 Décembre 2015

**ESSAI SUR L'INTEGRATION DES INTERACTIONS DANS
L'EVALUATION DES POLITIQUES PUBLIQUES : L'APPORT DES
METHODES DE L'ECONOMETRIE SPATIALE**

Directeur de thèse

Catherine BAUMONT

Membres du jury :

AYONG LE KAMA Alain, Professeur à l'Université Paris Ouest Nanterre La Défense,
Examineur

BAUMONT Catherine, Professeur à l'Université de Bourgogne, Directrice

L'HORTY Yannick, Professeur à l'Université Paris-Est Marne-La-Vallée, Rapporteur

LE GALLO Julie, Professeur à Agrosup Dijon, Examinatrice

GIVORD Pauline, Administrateur Hors Classe de l'INSEE, Examinatrice

MOIZEAU Fabien, Professeur à l'Université de Rennes 1, Rapporteur

La faculté n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses. Les opinions doivent être considérées comme propres à leurs auteurs

A mes feux grands-parents

A feu mon ami Adama N'Diath

Remerciements

A l'issue de la rédaction de cette recherche, je suis convaincu que la thèse est loin d'être un travail solitaire. La thèse a été un long parcours parsemé d'embûches, mais des Personnes et Amis ont toujours été à mes côtés pour me soutenir et m'encourager.

Je tiens tout d'abord à remercier ma directrice, Madame Catherine Baumont, sans qui ce rêve du petit écolier de *Golléré* (mon village natal) n'aurait pas pu se concrétiser. Je vous remercie pour m'avoir guidé, encouragé, conseillé durant ces années. Je tiens à vous exprimer, ici, par ces mots toute ma gratitude et toute ma reconnaissance.

Je tiens à remercier l'ensemble des membres du jury qui m'ont fait l'honneur d'accepter de participer à cette soutenance. Je remercie les professeurs Yannick l'Horty et Fabien Moizeau d'avoir accepté d'être les rapporteurs de cette thèse.

Je remercie également le PUCA qui par son soutien au programme de recherche « Approches économiques des dynamiques urbaines » m'a permis de réaliser les différentes applications de la thèse, et plus particulièrement Emmanuel Raoul, secrétaire permanent au PUCA, Lionel Martins, chargé de mission au PUCA et le professeur Alain Ayong le Kama, conseiller scientifique au PUCA. Je remercie notre gestionnaire maison, Jennifer Fourneret et Luc Carpentier, notre ingénieur d'étude pour son appui constant au combien important.

Je remercie l'ensemble du personnel de la MSH sans exception, une belle équipe disponible que j'ai eu le plaisir de côtoyer pendant trois belles années et en particulier Camille Grivault, mon voisin de bureau.

Mes remerciements vont à l'endroit de mes compagnons de route, avec qui j'ai tissé des relations qui vont au-delà d'une simple vie de labo : Docteur Benjamin Laurent, Aligui, Michael Flacandji (mes *compañeros* du début de l'aventure, salle 514 au Pôle d'Economie et de Gestion), Marion, Cyrenus, Yris, Daouda, Emmanuel, Fanny, Alexandre Chevet (*Docteur Honoris causa de la salle 514*), mes potes Alex Koumba, Momar Diop, Hamidou Diallo et Jo Camara.

Je remercie également le professeur Russel Davidson de l'Université de McGill de m'avoir accueilli pendant trois mois dans son laboratoire, pour l'hospitalité dont il a fait preuve envers moi lors de mon séjour.

Mes remerciements aux braves instituteur (e) s et professeur(e)s et maîtres de stage que j'ai côtoyés depuis mon école primaire jusqu'à l'Université de Bourgogne pour

le savoir, le savoir-être et le savoir-faire qu'ils m'ont donnés. Sans oublier ma tante Racky qui m'a toujours soutenu et encouragé dans mes études.

Je remercie également l'ensemble de mes amis au Sénégal particulièrement Amadou Diack, Abou Thiam, Mamadou Alpha Kane, pour leurs disponibilités, la distance ne nous a jamais séparés.

Enfin les remerciements ne seraient pas complets sans remercier ma chère épouse Diariata DIENG BA, ma *habibti* pour son soutien, sa patience et d'avoir toujours été à mes côtés surtout pendant les moments de doute. *A jaarama!*

A mon père El hadji Gory Ardo et à ma mère Hadja Bineta « *yo allah juunu balde* ». A mes frères Ardo, Ablaye, Mamoudou, Oumar, Amba , Welle et à ma petite sœur chérie Mairame. A Baba Mika Dieng et Athia pour leurs encouragements.

Table des matières

REMERCIEMENTS	V
TABLE DES MATIERES	VII
LISTE DES TABLEAUX	IX
LISTE DES FIGURES.....	X
PRESENTATION GENERALE	2
CHAPITRE 1 : INTEGRATION DES EFFETS D'INTERACTION DANS L'EVALUATION ECONOMETRIQUE DES EFFETS DE TRAITEMENT : ENJEUX ET METHODES.....	8
Introduction	9
1.1. Politiques publiques et contextes.....	12
1.2. Méthodes économétriques de l'évaluation des politiques publiques.....	16
1.3. Effets d'interaction et évaluation de l'effet de traitement	44
1.4. Intégration des effets d'interaction dans l'estimation de l'effet de traitement par l'économétrie spatiale	59
Conclusion	75
CHAPITRE 2 : LES INTERDEPENDANCES STRATEGIQUES DANS LES DECISIONS DE PARTICIPATION A UNE POLITIQUE TERRITORIALE	76
Introduction	77
2.1 Présentation et enjeux du projet Grand Paris.....	82
2.2 Modélisation de la décision de participation au Grand Paris	87
2.3 Présentation des données et des structures d'interactions spatiales	94
2.4 Résultats des estimations	102
Conclusion.....	108
CHAPITRE 3 : LA PRISE EN COMPTE DES INTERACTIONS DANS L'EVALUATION <i>EX ANTE</i> : UN MODELE DE DOUBLES DIFFERENCES SPATIAL APPLIQUE AU PROJET DU GRAND PARIS	110
Introduction	111
3.1 Présentation du projet du Grand Paris	116
3.2 Les interactions entre les collectivités : cadre théorique et empirique	123
3.3 Modélisation <i>Difference-in-Difference</i> du projet Grand Paris	128
3.4 Le modèle empirique : données et spécification des interactions	137
3.5 Résultats des estimations	149
Conclusion	158
CHAPITRE 4 : HETEROGENEITE DES INTERACTIONS ET EVALUATION DES EFFETS DE TRAITEMENT : UNE APPROCHE PAR LES EFFETS DE DEPENDANCE SPATIALE	162
Introduction	163
4.1. Interactions, interdépendances et évaluation des politiques publiques	167
4.2. Modèle d'évaluation avec interactions spatiales	175
4.3. Application empirique	187
Conclusion	204

CHAPITRE 5 : BOUNDING TREATMENT EFFECT AND INTERACTIONS: A SPATIAL ECONOMICS APPROACH APPLIED TO THE FRENCH “ZONES URBAINES SENSIBLES” PROGRAM	208
Introduction	209
5.1 Spatial modeling of nonparametric bounds on treatment effects	213
5.2 Interactions and identification of treatment effects	215
5.3 The French <i>Zones Urbaines Sensibles</i> (ZUS) program.....	224
5.4 Estimation methods and results	235
Conclusion.....	243
CONCLUSION GENERALE.....	245
BIBLIOGRAPHIE	249

Liste des tableaux

Tableau 1-1: Différents modèles pour l'estimation de la performance de la politique selon les types d'effets	56
Tableau 2-1: Description des matrices de voisinage	97
Tableau 2-2 : Statistiques descriptives des variables explicatives	101
Tableau 2-3 : Statistiques descriptives des variables explicatives (suite)	101
Tableau 2-4 : Résultats des estimations logistique standard et autologistic.....	106
Tableau 2-5 : Comparaison des capacités prédictives des modèles	107
Tableau 3-1 : Tableau récapitulatif des estimateurs en <i>DID et DID-spatial</i>	136
Tableau 3-2 : Statistiques descriptives des dépenses d'équipement.....	141
Tableau 3-3a : Statistiques descriptives des caractéristiques des communes.....	143
Tableau 3-4 : la composition des voisinages des communes	149
Tableau 3-5 : L'effet de traitement sans interactions entre les communes	150
Tableau 3-6 : Diagnostic de l'interdépendance spatiale.....	152
Tableau 3-7 : L'effet de traitement avec interactions entre les communes	154
Tableau 4-1 : Caractéristiques des communes selon leur statut	190
Tableau 4-2 : Variables exogènes du modèle d'évaluation.....	193
Tableau 4-3 : Typologies thématiques et répartition des communes	194
Tableau 4-4 : Caractéristiques des matrices de voisinage	195
Tableau 4-5 : Résultat de l'estimation de l'équation de sélection.....	197
Tableau 4-6 : Estimations des résultats pour les communes ZRR et hors ZRR.....	198
Tableau 4-7 : Estimations des effets de traitements	202
Table 5-1: The estimated bounds of ATE	219
Table 5-2: The estimated bounds of ATT	220
Table 5-3: ZUSs in Paris and its surrounding departments	227
Table 5-4: Logit regression for Propensity Score Matching	232
Table 5-5: Descriptive statistics for observable characteristics.....	234
Table 5-6: Description of neighborhood matrices	238
Table 5-7: Estimated bounds of the ZUS program effect on the unemployment rate under the worst case	244
Table 5-8: Estimated bounds of the ZUS program effect on the unemployment rate under the MTR ..	244

Liste des figures

Figure 1-1: Répartition de la population en sous population d'après les critères d'éligibilité.	15
Figure 1-2 : La décomposition de l'effet de causal (Halloran et Struchiner, 1991).....	58
Figure 1-3 : Les différentes formes d'interactions	60
Figure 1-4 : L'approche spatiale dans l'évaluation des politiques publiques.....	69
Figure 3-1 : Evolution des dépenses d'équipement des communes	139
Figure 4-1. Matrice de voisinage et typologie des interactions.....	180
Figure 4-2 : Matrice de voisinage et typologie des interactions.....	195
Figure 5-1: ZUS and IRIS overlap	230
Carte 2-1 : Les 7 pôles du Grand Paris.....	83
Carte 2-2 : Schéma d'ensemble du Grand Paris	84
Carte 2-3 : Situation des CDT – 1 ^{er} avril 2014.....	86
Carte 2-4 : Le périmètre d'étude et le projet « Grand Paris »	95
Carte 3-1 : Le tracé du métro Grand Paris Express et localisation des CDT	118
Carte 3-2 : Les territoires stratégiques du Grand Paris.....	121
Carte 4-1 : Répartition spatiale des communes ZRR et hors ZRR de l'étude.....	189
Map 5-1 : The Paris area: Paris and the <i>Petite Couronne</i>	226
Map 5-2: ZUS areas in Paris and its surrounding départements	228
Map 5-3 : the spatial distribution of IRISs in ZUS and Non ZUS groups	233

Présentation Générale

Cette thèse est composée de cinq chapitres indépendants mais liés par la question de la prise en compte des effets d'interaction dans l'évaluation d'une politique non expérimentale. Nous proposons une extension de quatre méthodes traditionnelles d'évaluation permettant d'intégrer les effets d'interaction dans l'évaluation d'une politique publique. Ces extensions s'appuient sur les méthodes de l'économétrie spatiale pour étudier et modéliser les effets d'interaction dans l'estimation de l'effet d'une politique publique.

Le premier chapitre introductif porte sur la problématique des effets d'interaction et l'évaluation des politiques non expérimentales. En effet, les méthodes d'estimation traditionnelles (*propensity score matching*, estimation en deux étapes ...) s'appuient sur des hypothèses excluant totalement (*stable unit treatment value assumption-SUTVA*) ou partiellement (*partial interference assumption-PIA*) la présence d'externalités liées au traitement. Or, dans le cadre d'une politique publique, les interactions entre les individus selon qu'ils sont bénéficiaires ou non, peuvent intervenir aussi bien dans le processus de sélection (pour les politiques non expérimentales surtout) que dans la performance de la politique. Par conséquent, la validité des hypothèses dans les méthodes traditionnelles est discutable. De plus, la non-prise en compte des effets d'interaction pourrait conduire à des estimations biaisées de l'effet de la politique et donc à de mauvaises conclusions sur l'effet réel de la politique. C'est pourquoi, nous proposons dans ce chapitre un cadre général permettant une extension des méthodes standards avec la prise en compte des effets d'interaction. Après une revue des principales méthodes traditionnelles d'évaluation et de leurs limites, nous abordons différents aspects des interactions, à savoir leurs sources principales, leurs natures et comment elles peuvent être intégrées dans la définition et l'estimation de l'effet d'une politique publique. Nous considérons différents cas : lorsque les interactions apparaissent (i) uniquement dans le processus de sélection (chapitre 2), (ii) uniquement dans la performance de la politique (chapitre 3 et 5) et (iii) à la fois dans le processus de sélection et dans la performance de la politique (chapitre 4).

Les quatre chapitres suivants sont alors consacrés à la prise en compte des interactions dans l'évaluation des politiques publiques, en proposant pour chaque chapitre, l'extension d'une méthode traditionnelle qui intègre les effets des

interdépendances dans l'évaluation des effets d'une politique. Chacun de ces chapitres fait l'objet d'une application empirique propre dont le choix a été raisonné afin de montrer la diversité des applications possibles, la thèse n'étant pas dédiée à l'évaluation d'une politique publique particulière. Dans chacun de ces chapitres, nous ne faisons donc pas une évaluation *stricto sensu* de la politique choisie à titre d'illustration.

Le deuxième chapitre est consacré exclusivement à la prise en compte des effets d'interaction dans le processus de sélection ou de participation des individus à une politique publique. Nous analysons les individus sous l'angle d'acteurs économiques dont les prises de décisions sont stratégiques, c'est-à-dire interdépendantes. Ainsi, chaque individu prend une décision optimale au sens économique, c'est-à-dire en tenant compte de la décision des autres individus. Pour tester empiriquement ces interdépendances stratégiques entre les acteurs, nous proposons une estimation d'un modèle auto-logistique, i.e. un modèle de choix binaire tenant compte des interdépendances des décisions individus. L'application empirique concerne le projet d'aménagement du Grand Paris et plus précisément le comportement des communes quant à leur décision de participer ou non à ce projet. La participation des communes prend la forme d'une adhésion à un Contrat de Développement Territorial, qui est une déclinaison locale – i.e. pour un ensemble de communes – des actions du Grand Paris en matière de logement, transport, emploi, environnement, culture, ... La modélisation du processus de décision par une spécification auto-logistique est déclinée avec deux schémas d'interdépendances (géographique et institutionnel). Nous montrons, selon le schéma considéré, qu'il existe des interdépendances stratégiques entre les communes avec un degré d'interdépendance stratégique beaucoup plus important dans le cas des interactions institutionnelles.

Le troisième chapitre propose une extension de la méthode *Difference-in-Difference* avec la prise en compte des interactions à travers une spécification autorégressive spatiale (modèle SAR). Avec cette modélisation, nous supposons que le traitement reçu par un individu peut affecter les résultats des autres individus de son voisinage.

L'application empirique porte sur l'estimation de l'effet du Projet Grand Paris sur les dépenses publiques par habitant des communes de l'Ile-de-France plus particulièrement concernées par le projet. A partir d'un modèle théorique de compétition

fiscale (Wildasin, 1988), nous supposons que le niveau de dépenses d'équipement des communes résulte à la fois de leurs propres caractéristiques mais aussi des interactions horizontales entre les communes. De ce fait, nous proposons la méthode *Difference-in-Difference Spatial* (SDID) pour tester l'effet du Grand Paris, en tenant compte des effets des interactions horizontales sur les dépenses d'équipement des communes. Nous considérons deux types de matrices de pondération pour capturer les effets de voisinage : une matrice économique, basée sur les flux domicile-travail et une matrice géographique, basée sur une distance seuil de 10 km. Nous estimons trois types d'effet du Grand Paris sur les dépenses des communes avec la méthode SDID sur trois périodes.

La première période 2001-2008 correspond à la période pré-annonce du projet en 2007. L'effet estimé correspond à l'effet d'anticipation sur l'annonce du Grand Paris. Le résultat montre que cet effet d'anticipation ainsi que l'effet d'interaction entre les communes sont non significatifs. La seconde période 2008-2012 correspond à la période post annonce du projet Grand Paris en 2007, consolidée par la loi relative à la constitution du Grand Paris et des CDT du 3 juin 2010. L'effet estimé correspond à l'effet d'annonce du Grand Paris. Les résultats montrent que dans un premier temps, cet effet, avec les deux matrices d'interactions sont positifs et significatifs au seuil de 10 %. Dans un second temps, l'effet d'interaction estimé est négatif et faiblement significatif avec la matrice des flux et largement significatif avec la matrice géographique. Quant à la troisième période 2001-2012, elle correspond à l'effet global de l'annonce du projet Grand Paris. C'est une période qui inclut les deux sous-périodes. Les résultats montrent que l'effet *ex ante* du projet Grand Paris et les effets d'interaction sont significatifs.

Ces résultats semblent indiquer que le projet Grand Paris a des effets positifs sur les dépenses d'équipement et qu'il existe des interactions horizontales entre les communes par rapport à leurs niveaux de dépenses.

Le quatrième chapitre est consacré à la prise en compte des effets d'interaction à la fois dans le processus de sélection et dans la performance de la politique. Nous nous intéressons au modèle d'estimation en deux étapes à la Heckman, lorsque la sélection est due à des inobservables. Nous proposons une version spatialisée de ce modèle qui permet d'obtenir des estimateurs des effets de traitement spatialement ajustés tenant compte des effets d'interaction. Lors de la première étape, nous modélisons un probit spatial de type *Spatial Error Model* qui permet de prendre en considération les effets d'interaction entre

les individus à travers leurs caractéristiques inobservables. Cette première étape permet de calculer l'Inverse de Ratio de Mills spatialisé. Lors de la seconde étape, nous supposons qu'il y a des effets d'interaction, c'est-à-dire que le traitement reçu par un individu peut affecter les résultats des autres individus de son voisinage. Nous estimons un modèle de type *Spatial Autoregressive* (SAR) qui permet d'intégrer ces effets de voisinage. Nous mettons en exergue dans ce chapitre l'existence des effets d'interaction aussi bien dans le processus de sélection que dans la performance de la politique sur une variable d'intérêt ou résultat observé. Le modèle prend en outre en compte l'hétérogénéité du voisinage, c'est-à-dire le fait qu'un individu a deux types de voisins - un voisinage composé des bénéficiaires de la politique et un autre composé d'individus non bénéficiaires. Nous modélisons cette hétérogénéité et estimons les effets différenciés des types de voisinage. Nous proposons une application portant sur l'effet de la classification des communes en Zones Rurales Revitalisées (ZRR) sur la création d'établissements. Nous montrons alors qu'il existe des effets de voisinage différenciés sur les créations d'établissements selon que les communes et leurs voisines sont classées ou non en ZRR. Nos résultats soulignent que la non prise en compte des effets d'interaction dans le processus de sélection, d'une part, et dans la performance de la politique, d'autre part, aurait conduit à une estimation biaisée de l'effet des ZRR.

Le cinquième et dernier chapitre est enfin consacré à l'intégration des effets d'interaction dans les méthodes non-paramétriques par bornes proposées par Manski (1989 ; 1990). Dans ces méthodes, il s'agit de ne pas proposer une estimation de la valeur-point, mais d'estimer des bornes de l'effet de traitement. En effet, cette approche constitue une alternative aux méthodes paramétriques qui s'appuient sur des hypothèses restrictives, comme l'hypothèse SUTVA, et semble appropriée lorsqu'on ne dispose pas de toutes les informations sur le processus de sélection des bénéficiaires. Cependant, la prise en compte des effets d'interaction constitue en soi une information supplémentaire et permet d'obtenir une meilleure estimation des bornes de l'effet de traitement. Le modèle que nous développons intègre les effets d'interaction en s'appuyant par ailleurs sur les extensions proposées dans les méthodes d'estimation par bornes de Manski et Pepper (2000). Pour le cas où il n'y a aucune hypothèse (« no assumption » ou « worst case ») sur les bornes estimées, ces dernières sont souvent trop larges et contiennent la valeur zéro ce qui ne permet pas de déterminer le signe de l'effet de la politique. Des

hypothèses supplémentaires sont alors introduites pour réduire l'étendue des intervalles et fournir le signe de l'effet du traitement. Nous proposons une application portant sur la politique de la ville et ses effets sur la redynamisation des quartiers prioritaires classés en Zones Urbaines Sensibles. Nous nous intéressons plus précisément aux quartiers prioritaires de Paris et des départements de la Petite Couronne. L'étude est menée à l'échelle des IRIS appartenant aux ZUS et nous estimons l'impact de la politique sur le taux de chômage. Nous montrons que lorsque des hypothèses sont faites sur la sélection (MTS) et les résultats (MTR), l'effet des ZUS est positif sur la réduction du chômage conditionnellement à certaines variables. L'estimation de l'effet de la ZUS est encore plus précise lorsqu'on considère à la fois les effets de voisinage et les effets spécifiques locaux de l'IRIS.

Chapitre 1 Intégration des effets d'interaction dans l'évaluation économétrique des effets de traitement : enjeux et méthodes

Introduction

Depuis les années 70, des méthodes statistiques et économétriques ont été développées pour déterminer l'effet causal de la politique, c'est-à-dire l'apport intrinsèque de la politique sur des changements intervenus à la suite de sa mise en place. Ces méthodes ont été introduites dans le domaine de la recherche médicale, mais ont été fortement utilisées, par la suite, dans l'évaluation des politiques de l'emploi et de la formation. Popularisées par les travaux de Rubin (1974), Rosenbaum et Rubin (1983), Holland (1986), Heckman (2001), Heckman et Vytlačil (2007), entre autres, ces méthodes ont permis la mise en place d'un cadre méthodologique pour estimer l'effet d'une politique publique.

Ces méthodes s'appuient sur des hypothèses fortes permettant d'identifier l'effet de traitement tout en contrôlant le problème du biais de sélection lié à l'identification des bénéficiaires. Cependant, une limite majeure de cette approche réside dans ces hypothèses, car elles sont restrictives, postulant une absence d'externalités de la politique donc excluant tout effet d'interaction entre les individus aussi bien dans le processus de sélection que dans les résultats observés.

Or, une politique publique est avant tout caractérisée par ses cibles et son périmètre d'action. Les individus ciblés, qu'ils soient bénéficiaires ou non, entretiennent des relations de voisinage entre eux qui conduisent à différentes formes d'effets d'interaction notamment appelées effets de voisinage, interactions sociales, effets de pairs.... De même, la sélection des bénéficiaires de la politique est généralement non aléatoire si bien que la décision d'un individu de participer ou non à la politique pourrait résulter de différentes formes d'interaction avec son voisinage.

Ces interactions peuvent être de nature sociale : la décision de participation d'un individu, en dehors de ses caractéristiques individuelles, peut être motivée par le comportement des autres individus de son groupe social, ethnique

Elles peuvent être de nature économique. Selon Manski (2000), la décision de participation d'un individu, en plus de ses caractéristiques intrinsèques, est motivée par différentes interactions. Les « *preference interactions* » influencent la décision prise par un individu en fonction des similitudes de préférences avec son voisinage. Les

« *expectations interactions* » amènent les individus à comparer les gains attendus avec les gains déjà obtenus par les individus qui ont déjà expérimentés la politique. Enfin, les « *constraint interactions* » désignent une décision prise de manière collective.

Les caractéristiques du périmètre d'action (quartier, commune, région,...) peuvent jouer un rôle important dans une politique publique notamment sur la diffusion de ses effets. La présence d'infrastructures de transport ou d'autres caractéristiques physiques de la zone d'action peut influencer la réussite d'une politique et la diffusion des effets sur les territoires voisins. A titre d'exemple, Briant *et al.* (2014, p.24) soulignent que « *deux dimensions géographiques ont notamment conditionné le succès du programme ZFU en France. L'accessibilité des transports, et plus particulièrement l'accès à la route et le train ou les stations de métro, améliore la capacité du programme à attirer des entreprises et créer des emplois. En revanche, les discontinuités générées par la circulation ou des barrières naturelles entre le quartier et les principaux centres d'emploi de son aire urbaine entravent les avantages d'allègements fiscaux et des exonérations sociales* ».

Ignorer ces différentes formes d'interaction entre territoires ou individus dans l'effet d'une politique peut conduire à une estimation biaisée de son effet réel : une sous-estimation si les effets d'interaction sont positifs et surestimation s'ils sont négatifs. Par ailleurs du point de vue économétrique, la non prise en compte de ces effets peut être assimilée à un problème d'endogénéité, c'est-à-dire un problème de variables omises- de simultanéité ou d'erreurs de mesure (Le Gallo et Mutl, 2014 ; Soetevent, 2006) et par conséquent les méthodes standards pour estimer l'effet de traitement ne sont plus adaptées.

L'objectif de ce chapitre est de proposer une extension des méthodes standards de l'évaluation en intégrant les effets des interactions dans l'estimation de l'effet de traitement. Cette approche vise en premier lieu à redéfinir l'effet de traitement tenant compte des effets d'interaction aussi bien dans le processus de sélection des bénéficiaires que dans les résultats observés. Nous proposons en deuxième lieu des méthodes d'estimation qui permettent de modéliser et d'estimer ces effets d'interaction et nous nous appuyons pour cela sur les méthodes de l'économétrie spatiale.

Le chapitre est organisé en cinq sections. La première section (1.1) rappelle le cadre des politiques publiques, les méthodes de ciblage des bénéficiaires et les erreurs liées à la répartition de la population. Dans la deuxième section (1.2), nous présentons l'effet de traitement, les paramètres d'intérêt et les principales méthodes économétriques d'évaluation utilisées. Nous terminons cette partie en insistant sur les limites à la fois économétriques et économiques de l'absence des effets d'interaction dans les estimations. La troisième section (1.3) est alors exclusivement dédiée aux effets des interactions. Elle porte sur les différentes sources d'interactions dans le cadre d'une politique publique et sur les mécanismes de diffusion des interactions. Enfin, dans la quatrième section (1.4), nous proposons un cadre méthodologique pour intégrer les effets d'interaction dans l'estimation de l'effet de la politique. Nous proposons une extension spatiale des paramètres d'intérêt en intégrant les effets d'interaction à la fois dans le processus de sélection et dans les résultats observés. Ces extensions portent sur les méthodes d'estimation de l'effet de traitement lorsque la sélection est faite sur des observables (méthodes *de matching*, *Difference-in-Difference*,...), les méthodes d'estimation lorsque la sélection est due aux inobservables (estimation en deux étapes à la Heckman) et les estimations non-paramétriques de l'effet de traitement à la Manski (1990).

1.1. Politiques publiques et contextes

Une politique publique est un ensemble d'actions visant à améliorer une situation, économique, sociale, politique ... donnée et peut concerner une partie de la population ou la population dans son ensemble. Une politique publique est caractérisée d'une part, par sa pertinence, c'est-à-dire la cohérence entre les problèmes à résoudre et les solutions proposées et d'autre part, par sa portée, c'est-à-dire l'étendue de la politique en termes de cibles (individus, territoires) ou en termes d'impacts (les effets induits dans le temps et dans l'espace).

Les politiques peuvent être classées en deux types selon la méthode de ciblage ou d'identification des bénéficiaires. Le premier concerne les politiques expérimentales, expériences contrôlées ou « *social experiment* », caractérisées par un protocole d'identification aléatoire des cibles. Dans ce type de politique, la répartition des individus (ménages, communes, entreprises, individus) dans le groupe des bénéficiaires et des non bénéficiaires est faite de manière aléatoire. Le cas des 10 000 permis de la réussite par exemple.

Le second type concerne les politiques non expérimentales qui se distinguent des politiques expérimentales, non pas tant sur la sélection des individus, mais sur le processus de sélection de ces individus qui n'est plus aléatoire. Elles sont caractérisées par la définition de critères d'éligibilité, qui conduiront à la répartition de la population entre les bénéficiaires et les non bénéficiaires sans que le processus aléatoire de cette répartition ait été contrôlé. Les critères d'éligibilité sont généralement définis par les décideurs publics et les différents acteurs impliqués dans la politique en question.

Les politiques publiques peuvent être caractérisées par les critères de sélection des bénéficiaires ou cibles. La sélection des bénéficiaires peut être tout d'abord fondée sur la satisfaction de critères précis associés à des indicateurs économiques et sociaux. Ces critères peuvent porter sur un revenu seuil des ménages pour les politiques sociales (Revenu de Solidarité Active, par exemple), le nombre d'enfants, le type de logement (politique de rénovation urbaine)... Ce type de politique est restrictif et destiné uniquement aux individus remplissant les critères de sélection.

Le ciblage des bénéficiaires peut être effectué sur diverses catégories : la taille des entreprises (pour une politique d'exonération fiscale), les jeunes de 16 à 18 ans pour les

politiques de santé jeunes (programme *Pass santé jeune*).... Dans ce type de politique, tous les individus appartenant à la catégorie sont éligibles.

Les politiques peuvent être géographiquement ciblées et on parle de politiques territorialisées. Le ou les critères de ciblage sont de nature géographique (village, commune, quartier, région, proximité d'une route, centrale électrique,...) ou institutionnel (appartenance au canton, au département ...). Ainsi, l'ensemble des individus localisés ou appartenant à l'espace géographique sont éligibles à la politique. Ce type de politique peut être assorti de critères d'éligibilité pour le territoire concerné : par exemple, les politiques territorialisées comme les zones rurales revitalisées (ZRR), les politiques urbaines dédiées aux quartiers en difficultés (zones urbaines sensibles (ZUS), les zones franches urbaines (ZFU), les zones d'éducation prioritaire (ZEP),...), les politiques européennes (FEDER)...

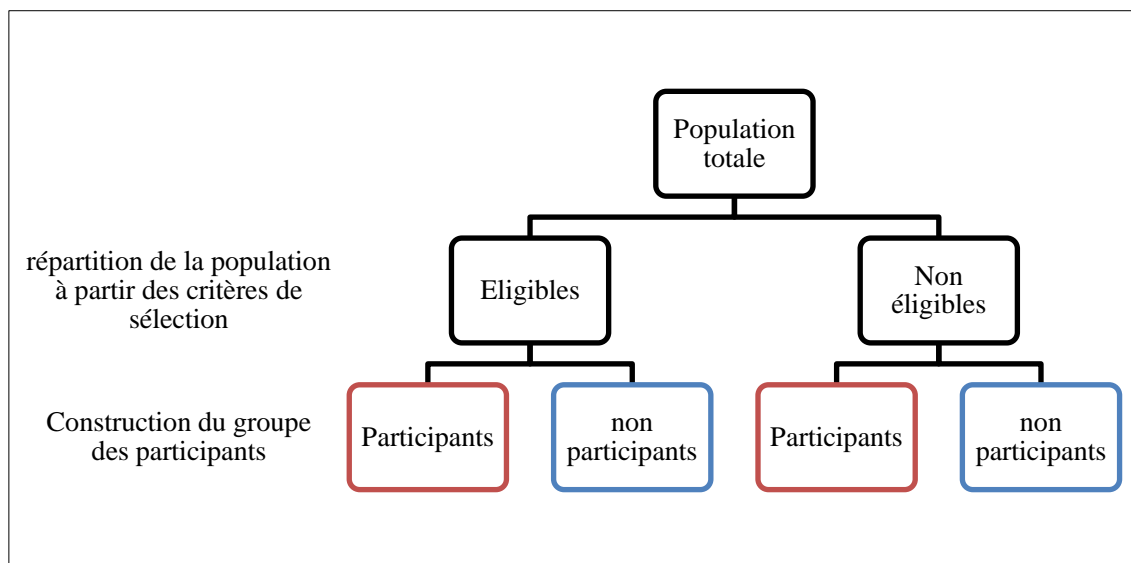
Les politiques « universelles » désignent enfin des politiques où l'ensemble de la population, sans restriction, est éligible. C'est une politique sans critères de sélection et par conséquent accessible à tous. L'ensemble des bénéficiaires de cette politique est constitué par des individus qui ne peuvent s'y soustraire (comme dans les politiques de santé publique) ou qui y participent de manière volontaire : par exemple la politique « travail contre nourriture » du Programme Alimentaire Mondial dans les pays en développement.

Ces règles de sélection classent la population totale d'une unité géographique ou d'un groupe en deux sous populations. D'une part, une sous population composée d'individus éligibles, c'est-à-dire l'ensemble des individus qui remplissent les critères d'éligibilité à la politique et de l'autre côté une sous population composée d'individus non éligibles, c'est-à-dire ceux qui ne remplissent pas les critères d'éligibilité (*cf.* Figure 1-1). Dans le cas des politiques expérimentales, les individus composant le second groupe peuvent satisfaire les critères, mais ils ne sont pas aléatoirement choisis pour faire partie des bénéficiaires. C'est le cas par exemple de la politique 10 000 permis pour la réussite où les individus ont été aléatoirement répartis entre les deux groupes selon le protocole. Chaque individu ayant donc connaissance du protocole et sachant qu'il pourra au final être bénéficiaire ou non et acceptant d'être observé tout au long de la mise en œuvre de la politique.

Au-delà du seul fait d'avoir le droit de bénéficier d'une politique publique, cela ne signifie pas que l'individu éligible fera partie de la politique. L'éligibilité ne signifie pas forcément participation. Certains individus éligibles, pour plusieurs raisons généralement inobservables par le chercheur, peuvent décider de ne pas participer à la politique. C'est pourquoi, la population éligible peut être subdivisée en participants et non participants (*cf.* Figure 1-1).

Par ailleurs, certaines populations non éligibles, compte tenu du niveau d'information dont elles disposent sur les politiques, peuvent également ajuster certaines de leurs caractéristiques pour bénéficier de la politique. C'est le cas de certaines politiques considérées comme stigmatisantes où des individus éligibles peuvent se soustraire à la participation par crainte d'être stigmatisés par les autres membres de la société. Nous pouvons citer en exemple les politiques d'assistance sociale ainsi que certaines politiques de santé (dépistage SIDA). C'est également le cas des politiques où la participation fait perdre certains avantages. Dans d'autres cas, les individus non éligibles, dont les caractéristiques sont très proches des seuils d'éligibilité, peuvent ajuster leurs comportements afin de bénéficier de la politique.

Figure 1-1: Répartition de la population en sous population d'après les critères d'éligibilité.



Source : Sileymane Ba-2015

Outre les caractéristiques de l'individu, la décision de participation d'un individu peut être influencée soit par la décision des autres individus de son voisinage dès lors que cet individu appartient à des groupes ou réseaux avec lesquels il entretient de fortes relations, soit par les caractéristiques de son voisinage.

Ces méthodes de ciblage, contrairement aux expériences contrôlées, sont ainsi fortement susceptibles d'induire des erreurs dans la répartition de la population entre les bénéficiaires et les non-bénéficiaires. Généralement, ces erreurs sont de deux types, à savoir les erreurs d'inclusion et les erreurs d'exclusion. Les premières se produisent lorsque les individus non éligibles sont inclus dans les bénéficiaires alors que les erreurs d'exclusion se produisent lorsqu'on exclut du groupe des bénéficiaires des individus, pourtant éligibles. En conséquence, si ces erreurs ne sont pas contrôlées, elles peuvent conduire à une évaluation biaisée de l'effet de la politique. Nous reviendrons plus en détails sur les sources et les conséquences de ces erreurs tout au long de ce chapitre.

Evaluer l'effet d'une politique, c'est procéder à une analyse de la performance de celle-ci sur une variable d'intérêt (revenu, réussite scolaire, taux de chômage, ...) sur un ensemble d'individus (au sens large). De manière plus large, si l'évaluation d'une politique peut être axée sur plusieurs indicateurs notamment l'efficacité, l'efficience, les

effets et l'impact, nous nous intéressons uniquement à l'évaluation de l'effet de la politique ou encore à l'effet causal de la politique. Il s'agit d'estimer l'apport « intrinsèque » de la politique sur les changements attendus à travers les méthodes économétriques de l'évaluation (1.2)

1.2. Méthodes économétriques de l'évaluation des politiques publiques

La terminologie utilisée dans l'évaluation économétrique provient de la médecine et de la biologie. C'est pourquoi, dans la suite, nous utilisons le terme "traitement" pour désigner la politique qui est un traitement appliqué à des individus (bénéficiaires) et qui détermine en même temps les individus non bénéficiaires. Ces derniers sont alors qualifiés d'individus de contrôle, car leurs caractéristiques observables sont similaires à celles des bénéficiaires. Après avoir rappelé et défini l'effet de traitement (1.2.1), nous présenterons les différentes méthodes d'évaluation qui seront ensuite mobilisées dans la thèse. Les méthodes quasi-expérimentales (1.2.2) : les méthodes d'appariement ou *matching*, les méthodes de double différences, adaptées lorsqu'on suppose que le biais de sélection est dû aux observables. La méthode d'estimation à la Heckman, lorsque le biais de sélection est dû aux inobservables (1.2.3) et enfin les approches non-paramétriques à la Manski « *bounding approach* » (1.2.4).

1.2.1 Définition de l'effet de traitement

En évaluation économétrique, l'effet de traitement fait référence aux effets induits et attribuables à la politique. Plusieurs méthodes ont été développées pour estimer cet effet, dont la plus utilisée, le modèle canonique de l'évaluation par les effets de traitement (Rubin, 1974). Elle s'appuie sur les méthodes statistiques et économétriques pour définir l'effet causal de la politique à évaluer tout en contrôlant les potentiels biais qui peuvent altérer la qualité de l'estimation et par conséquent interroger les implications socio-économiques des politiques publiques.

Dans le cadre du modèle causal de Rubin (Rubin, 1974) noté MCR, chaque individu est caractérisé par deux résultats potentiels, y_1 et y_0 , respectivement les résultats avec traitement et sans traitement.

Soit y_{ik} , le résultat observé d'un individu i selon son statut k avec $k = 1$, si l'individu est bénéficiaire et $k = 0$, s'il est non bénéficiaire. On suppose que le résultat observé y_{ik} est une fonction linéaire des caractéristiques observables et inobservables.

$$y_{ik} = x' \beta_k + \varepsilon_i ; k = \{0,1\} \quad (1-1)$$

Avec x un vecteur $(1 \times p)$ des variables explicatives, ε_i représente les caractéristiques inobservables suivant une loi normale et β_k , le paramètre associé aux variables explicatives x à estimer.

L'effet de la politique sur un individu i est ainsi défini comme étant la différence entre son résultat observé sous traitement et son résultat observé sans le traitement.

$$\alpha_i = y_{1i} - y_{0i} \Leftrightarrow y_{1i} = y_{0i} + \alpha_i \quad (1-2)$$

Cependant, l'effet causal α_i tel que défini est concrètement inobservable et il ne peut pas être déterminé à partir de ces deux résultats potentiels, car ils ne peuvent pas être simultanément observés pour chaque individu i . Pour un individu traité on observe uniquement son résultat potentiel y_{1i} et son résultat potentiel y_{0i} est inobservable. La variable y_{0i} est appelée *contrefactuel* et correspond au résultat qui aurait été observé si l'individu n'avait pas été traité. Pour un individu non traité on observe uniquement son résultat potentiel y_{0i} et son *contrefactuel* y_{1i} est inobservable.

Ainsi, le résultat observé pour un individu i (y_i) est donné par la relation entre les résultats potentiels et la variable de traitement D , qui est égale à 1 si l'individu est traité et zéro s'il n'est pas traité.

$$y_i = D y_{1i} + (1 - D) y_{0i} \quad (1-3)$$

Par conséquent, pour identifier l'effet de la politique, il convient donc de définir des méthodes pour disposer de ces informations nécessaires à l'identification de l'effet du traitement.

Par ailleurs, compte tenu du nombre de bénéficiaires, il est plus pertinent de s'intéresser à l'effet de la politique sur un ensemble d'individu que de déterminer l'effet du traitement sur un individu. C'est pourquoi dans la littérature, on s'intéresse généralement à trois types d'effet causal appelés paramètres d'intérêt à l'échelle de la population.

1 - L'effet moyen du traitement dans la population, noté ATE, mesure l'impact de la politique sur l'ensemble de la population, c'est-à-dire l'effet de la politique sur les traités par rapport aux non traités. Il est défini comme étant la différence des résultats moyens observés dans le groupe des traités et dans le groupe de contrôle :

$$\begin{aligned}ATE &= E(\alpha) = E(Y_1 - Y_0) \\ &= E(Y_1|D = 1) - E(Y_0|D = 0)\end{aligned}\tag{1-4}$$

2 - L'effet moyen du traitement sur les traités, noté ATT, mesure l'impact de la politique uniquement sur les traités. Il est défini comme étant la différence de leurs résultats moyens observés sous traitement et leurs résultats moyens observés sans le traitement :

$$\begin{aligned}ATT &= E(\alpha|D = 1) \\ &= E(Y_1 - Y_0|D = 1) \\ &= E(Y_1|D = 1) - E(Y_0|D = 1)\end{aligned}\tag{1-5}$$

3 - L'effet de traitement moyen sur les non traités, noté ATU, mesure l'impact que la politique aurait eu sur les non bénéficiaires. Il est défini comme étant la différence entre leurs résultats moyens observés s'ils avaient été sous traitement et leurs résultats moyens observés sans le traitement :

$$\begin{aligned} ATU &= E(\alpha|D = 0) \\ &= E(Y_1 - Y_0|D = 0) \\ &= E(Y_1|D = 0) - E(Y_0|D = 0) \end{aligned} \tag{1-6}$$

L'effet de traitement total ATE peut être déduit à partir des équations (1-5) et (1-6) :

$$ATE = \frac{N_1}{N} ATT + \frac{N_0}{N} ATU \tag{1-7}$$

où N représente la population totale, N_1 la population des individus traités et N_0 la population des individus de contrôle.

A partir de ces définitions, nous constatons que le paramètre ATE peut être déterminé directement à partir des données observées. En revanche, les paramètres ATT et ATU ne sont pas observables, car ils dépendent respectivement des *contrefactuels* $[E(Y_0|D = 1)]$ et $[E(Y_0|D = 0)]$.

Pour estimer l'effet moyen de traitement sur les traités, ATT , les résultats observés sur des individus de contrôle sont utilisés en postulant l'hypothèse suivante : on suppose que les résultats des individus de contrôle sont proches de ceux que les individus traités auraient connus s'ils n'avaient pas été traités, c'est-à-dire :

$$E(Y_0|D = 0) \cong E(Y_0|D = 1) \tag{1-8}$$

On définit un nouveau paramètre ΔATT qui est l'estimateur de l'effet moyen de traitement sur les traités ATT .

$$\Delta ATT = E(Y_1|D = 1) - E(Y_0|D = 0) \tag{1-9}$$

En réécrivant l'équation (1-9) on obtient une nouvelle écriture de ΔATT :

$$\begin{aligned}\Delta ATT &= E(Y_1|D = 1) - E(Y_0|D = 1) + E(Y_0|D = 1) - E(Y_0|D = 0) \\ &= ATT + E(Y_0|D = 1) - E(Y_0|D = 0) = ATT + BS\end{aligned}\tag{1-10}$$

Le second terme de l'équation (1-10), noté BS, représente le biais de sélection. Il est présent lorsque les bénéficiaires de la politique ne sont pas représentatifs de l'ensemble de la population éligible (L'Horty et Petit, 2011) et résulte principalement des erreurs d'inclusion ou d'exclusion. Il faut noter que, contrairement aux politiques non expérimentales, les expériences contrôlées ne sont pas exposées au problème du biais de sélection d'où leur appellation « *the gold standard* »¹ dans l'évaluation des politiques.

Il existe plusieurs types de biais de sélection (Maddala, 1983). *Le biais d'auto-sélection* est très présent dans les politiques à participation volontaire et dans le cadre des politiques « stigmatisantes ». Ce biais fait référence au fait que les individus, compte tenu de leurs caractéristiques non observables ou « attributs ou habilités » (Goldberger, 2008) vont « s'auto-sélectionner ». Par exemple, dans une politique à participation volontaire, les individus qui ont des avantages comparatifs positifs par rapport à la politique seront plus susceptibles de participer à la politique. Dans ce cas, la politique aura un effet important, mais surestimé par rapport à une expérience aléatoire. En revanche, dans le cas des politiques stigmatisantes, certains individus, pourtant éligibles auront tendance à s'autocensurer conduisant à une erreur d'exclusion.

Le biais de sélection de l'échantillon fait référence au fait que la variable dépendante n'est observée que pour les individus sélectionnés qu'ils soient dans le groupe de contrôle ou dans le groupe des traités.

Le biais de sélection lié aux gestionnaires du programme, en revanche, est la résultante d'une sélection par les gestionnaires de telle sorte que le « groupe de traitement » serait composé de personnes relativement moins défavorisées que le « groupe témoin » (Barnow et al, 1980; Goldberger, 2008).

¹ Heckman et Vytlačil (2007, p.4881)

Le biais de sélection est alors défini comme étant la différence entre l'estimateur ΔATT (équation 1-10) et ATT (équation 1-5).

$$BS = [E(Y_0|D = 1) - E(Y_0|D = 0)] \quad (1-11)$$

Si le biais de sélection est nul, alors :

$$[E(Y_0|D = 1) - E(Y_0|D = 0)] = 0 \Rightarrow E(Y_0|D = 1) = E(Y_0|D = 0)$$

Dans ce cas, l'estimateur ΔATT est égal à l'effet moyen de traitement dans la population et est tout simplement égal à la différence des résultats moyens observés pour les bénéficiaires et les non bénéficiaires.

$$\Delta ATT = ATT = ATE = E(Y_1|D = 1) - E(Y_0|D = 0) \quad (1-12)$$

En revanche si le BS est non nul, alors l'estimateur de l'effet moyen du traitement sur les traités est composé de l'effet propre de la politique ATT et du biais de sélection.

Ce dernier est susceptible d'être systématiquement présent dans les politiques publiques non expérimentales. Les paramètres d'intérêt sont alors fortement exposés à ce problème, car le résultat moyen du groupe de contrôle ne peut pas coïncider exactement avec celui du contrefactuel quand $BS \neq 0$. En ignorant le biais de sélection, on obtient un estimateur de l'effet de traitement biaisé et avec deux conséquences principales.

Premier cas : si le biais de sélection est positif alors l'estimateur Δ est surestimé

$$\Delta ATT = ATT + BS > ATT \quad (1-13)$$

Deuxième cas : si le biais de sélection est négatif alors l'estimateur Δ est sous-estimé

$$\Delta ATT = ATT + BS < ATT \quad (1-14)$$

Pour obtenir des estimateurs non biaisés des paramètres d'intérêt, différentes méthodes d'estimation ont été développées pour contrôler le biais de sélection. D'une part, nous avons des méthodes d'estimation lorsque le biais de sélection est dû aux observables, c'est-à-dire à la non prise en compte des différences sur les observables entre le groupe de traités et le groupe de contrôle. D'autre part, des méthodes d'estimation sont développées lorsque le biais de sélection est dû aux inobservables, c'est-à-dire à la non prise en compte des différences sur les inobservables des deux groupes. Par ailleurs, en considérant que le biais de sélection est dû aux observables, cela signifie que le biais dû aux inobservables est ignoré.

Dans la partie suivante, nous présentons les principales méthodes dont certaines seront utilisées ensuite dans la thèse tout en sachant qu'il existe d'autres méthodes d'estimation (méthode des variables instrumentales, la méthode de *Regression Discontinuity Design*,...).

Les méthodes d'estimation des paramètres d'intérêt peuvent être classées en trois groupes. (i) Les méthodes quasi-expérimentales : les méthodes d'appariement ou *matching*, les méthodes de double différences, adaptées lorsqu'on suppose que le biais de sélection est dû aux observables. (ii) La méthode d'estimation à la Heckman, lorsque le biais de sélection est dû aux inobservables. Et enfin (iii) les approches non-paramétriques à la Manski « *bounding approach* ».

1.2.2 Les méthodes quasi expérimentales d'appariement

Le principe général de ces méthodes quasi expérimentales porte sur la constitution du groupe de contrôle. Il s'agit de définir les conditions d'appariements (*matching*) entre l'individu traité et un individu jumeau non bénéficiaire de la politique.

1.2.2.1 Principes et hypothèses

Le *matching* ou l'appariement dans le cadre de l'évaluation de l'effet de traitement consiste à trouver pour chaque individu traité un ou plusieurs individus de contrôle dont les caractéristiques observables sont similaires. Comme le souligne Fougère (2000, p.112), « *c'est la méthode qui se rapproche le plus des expérimentations contrôlées parce que la mise en œuvre de ces deux méthodes (expérimentation et appariement) ne dépend*

pas directement de la valeur prise par les variables de résultat (...) permettant d'évaluer l'efficacité du dispositif ».

L'identification des paramètres d'intérêt avec les méthodes *matching* requiert trois hypothèses fortes :

Hypothèse I² : l'hypothèse de l'indépendance conditionnelle (CIA pour *Conditional Independence Assumption*). On suppose que les résultats potentiels sont indépendants au traitement conditionnellement à des caractéristiques observables X, non affectées par le traitement. Formellement :

$$(Y_1, Y_0) \perp D | X \Leftrightarrow E(Y_0 | D = 1, X) = E(Y_0 | D = 0, X) \quad (1-15)$$

Cette hypothèse implique que la sélection des bénéficiaires est uniquement basée sur les caractéristiques observables et que toutes les variables qui agissent simultanément sur la sélection des bénéficiaires et sur les résultats potentiels sont également observées par le chercheur.

En outre, en contrôlant les caractéristiques observables X, cette hypothèse implique que la répartition de la population peut être considérée, dans ce cas, comme résultant d'une affectation aléatoire (Brodsky et al., 2007), conduisant ainsi à deux groupes comparables – un groupe de traitement et un groupe de contrôle. Cela signifie alors que la distribution du *contrefactuel* ($Y_0 | D = 1$) est la même que la distribution du résultat observé pour le groupe de contrôle ($Y_0 | D = 0$), avec comme corollaire la nullité du biais de sélection $BS = 0$ (équation 1-10).

Cependant, lorsque le *matching* sur les observable nécessite de nombreuses variables, l'appariement entre un individu traité et un individu de contrôle devient compliqué, car il est difficile de trouver deux individus similaires dans les deux groupes.

² Cette hypothèse est connue aussi dans la littérature, sous différentes appellations notamment « *ignorable treatment assignment* » (Rosenbaum and Rubin, 1983a) ; « *selection on observables* » « *fundamental identification problem* » (Barnow, Cain, et Goldberger, 1980) ou encore « *unconfoundedness* » (Imbens et Wooldridge, 2009).

Pour contourner cette limite, Rosenbaum et Rubin (1983) proposent l'hypothèse d'indépendance conditionnelle (CIA) basée sur le score de propension.

Le score de propension noté PS , est défini comme étant la probabilité pour un individu de participer à un traitement, conditionnellement à ses caractéristiques observables X .

$$PS = P(X) = P(D = 1|X) \quad (1-16)$$

Les auteurs montrent que si les résultats potentiels sont indépendants du traitement conditionnellement aux caractéristiques observables X , ils sont également indépendants du traitement conditionnellement à un score de propension. L'hypothèse *CIA*, basée sur le score de propension, s'écrit de la manière suivante :

$$(Y_1, Y_0) \perp D | P(X) \Leftrightarrow E(Y_0|D = 1, P(X)) = E(Y_0|D = 0, P(X)) \quad (1-17)$$

Hypothèse 2 : l'hypothèse de support commun ou overlap assumption

On suppose que pour les individus traités, il existe des individus de contrôle ayant des caractéristiques similaires. L'hypothèse de support commun signifie que les individus dont les caractéristiques sont similaires, ont une probabilité positive d'être soit des participants soit des non-participants (Heckman et *al.*, 1999, p.55). Formellement :

$$0 < Pr(D = 1|X = x) < 1 \text{ pour tout } x \quad (1-18)$$

Elle a comme corollaire l'annulation du biais de sélection, car conditionnellement au score de propension, on a :

$$E(Y_0|D = 1, P(X)) = E(Y_0|D = 0, P(X)) \Rightarrow ((E(Y_0|D = 1) - E(Y_0|D = 0)) = 0)$$

Hypothèse 3 : l'hypothèse Stable Unit Treatment Value Assumption (SUTVA).

Par cette hypothèse, on suppose que le traitement n'affecte que le résultat de l'individu traité. Cela signifie qu'il y a absence d'interférences entre les individus donc une absence d'externalités du traitement.

1.2.2.2 La méthode de *matching* par le *propensity score matching* (PSM)

L'estimateur du PSM correspond à la moyenne de la différence des résultats sur le support commun, pondérée par la distribution du score de propension des traités. L'estimation de l'effet de traitement dans ce cadre se déroule en quatre étapes.

Dans la première étape, une équation de sélection est fonction d'un certain nombre de variables explicatives qui affectent la probabilité d'être sélectionné. Les variables explicatives à intégrer dans le modèle de sélection sont l'ensemble des variables qui affectent à la fois le processus d'affectation au traitement et le résultat. En deuxième étape, on estime l'équation de sélection à travers un modèle de choix binaire (probit ou logit) et on calcule les probabilités estimées. Dans la troisième étape, on choisit un critère de similarité (plus proches voisins, l'estimateur à noyau ou la distance seuil (*caliper*), pour faire l'appariement à partir des probabilités estimées. Enfin, la dernière étape consiste à estimer les paramètres d'intérêt avec les échantillons appariés.

Les estimateurs des paramètres d'intérêt sont donnés par les équations (1-19), (1-20) et (1-21) :

L'estimateur de l'effet de traitement moyen sur les traités (équation (1-5) est :

$$\Delta ATT = \frac{1}{N_1} \sum_{i \in \{D_i=1\}} \left(y_{1i} - \frac{1}{\sum_{j \in \{D_j=0\}} \omega_{ij}} \sum_{j \in \{D_j=0\}} \omega_{ij} y_{0j} \right) \quad (1-19)$$

L'estimateur de l'effet de traitement moyen sur les non traités (équation (1-6) est :

$$\Delta ATU = \frac{1}{N_0} \sum_{i \in \{D_i=0\}} \left(\frac{1}{\sum_{j \in \{D_j=1\}} \omega_{ij}} \sum_{j \in \{D_j=1\}} \omega_{ij} y_{1j} - y_{0i} \right) \quad (1-20)$$

L'estimateur de l'effet moyen de traitement total (équation (1-4) est :

$$\Delta ATE = \frac{N_1}{N} \Delta ATT + \frac{N_0}{N} \Delta ATT \quad (1-21)$$

Avec $N_j = \sum_i \mathbb{1}[D_i = j]$, $j = \{0,1\}$, la taille de chaque groupe.

Les variables $\{y_{1i}, y_{1j}; y_{0i}, y_{0j}\}$ représentent les résultats observés des individus dans le groupe traité et le groupe de contrôle. Les termes ω_{ij} représentent les pondérations entre les individus i et j à définir.

Il existe plusieurs possibilités pour spécifier la pondération appropriée à associer entre un individu traité et son « jumeau » dans le groupe de contrôle. La pondération ω_{ij} est fonction d'une « distance » entre les deux individus, et plus cette distance est faible, plus les deux individus sont proches ou similaires.

Dans le cas d'un *matching* avec la méthode PSM, la distance entre les deux individus est définie comme étant les différences dans les probabilités d'être traités ou non conditionnellement à X. Cette distance est donnée par :

$$d_{ij} = |P(x_i) - P(x_j)| \quad (1-22)$$

En utilisant les scores de propension, on définit $C(P(x_i))$ comme étant le voisinage de l'individu i autour de la probabilité $P(x_i)$ et A_i l'ensemble des individus de contrôle appariés à l'individu i autour de ce voisinage $C(P(x_i))$, avec :

$$A_i = \{j | D_j \neq D_i, P(x_j) \in C(P(x_i))\} \quad (1-23)$$

Pour déterminer la pondération ω_{ij} plusieurs techniques existent et nous en exposons les plus courantes.

- **Le critère du plus proche voisin**

Le voisinage $C(P(x_i))$ est défini par la minimisation de la distance d_{ij} :

$$\begin{aligned} C(P(x_i)) &= \min_j |d_{ij}| \\ \Rightarrow \omega_{ij} &= \begin{cases} 1, & \text{si } j \in A_i \\ 0, & \text{sinon} \end{cases} \end{aligned} \quad (1-24)$$

L'individu j voisin de l'individu i correspond à l'individu du groupe de contrôle ayant le score de propension estimé le plus proche à celui de l'individu i .

- **Le critère des k -plus proches voisins**

Dans ce cas, les individus j voisins de l'individu i correspondent aux k individus du groupe de contrôle ayant les scores de propension les plus proches de celui de l'individu i autrement les k -plus proches voisins de l'individu i . Le voisinage $C(P(x_i))$ est défini par la minimisation des k -distance d_{ij} :

$$\begin{aligned} C(P(x_i)) &= k - \min_j |d_{ij}| \\ \Rightarrow \omega_{ij} &= \begin{cases} 1/k, & \text{si } j \in A_i \\ 0, & \text{sinon} \end{cases} \end{aligned} \quad (1-25)$$

- **Caliper or radius matching**

Dans ce cas, le voisinage est défini à partir d'une valeur seuil ε du score de propension estimé. Les individus j voisins de l'individu i correspondent à l'ensemble des individus du groupe de contrôle dont le score de propension est inférieur au seuil ε .

$$\begin{aligned} C(P(x_i)) &= \{P(x_j) | d_{ij} < \varepsilon\} \\ \Rightarrow \omega_{ij} &= \begin{cases} 1/k_i, & \text{si } j \in A_i \\ 0, & \text{sinon} \end{cases} \end{aligned} \quad (1-26)$$

- **Estimateur à noyau ou *Kernel matching***

Chaque individu traité est apparié avec la moyenne pondérée des individus de contrôle où les poids sont inversement proportionnels à la distance entre les scores de propension du groupe des traités et du groupe de contrôle.

$$C(P(x_i)) = \left\{ \left| \frac{P(x_j) - P(x_i)}{h} \right| \leq \varepsilon \right\} \quad (1-27)$$

$$\Rightarrow \omega_{ij} = \begin{cases} \varphi, & \text{si } j \in A_i \\ 0, & \text{sinon} \end{cases}$$

Avec $\varphi = \frac{K\left|\frac{P(x_j)-P(x_i)}{h}\right|}{\sum_{j' \in \{D_{j'}=0\}} K\left|\frac{P(x_{j'})-P(x_i)}{h}\right|}$; $K(\cdot)$ est la fonction de noyau et h la fenêtre.

1.2.3 Les méthodes de régression

Pour estimer les paramètres d'intérêt avec les méthodes de régression, nous considérons, à titre d'exposition, un cadre de régression des données de panel (Todd, 2008) qui montre seulement que les méthodes économétriques d'évaluation sont développées également dans ce contexte.

Soit Y_{1it} et Y_{0it} les équations de résultats de l'individu i à la période t , où :

$$y_{1it} = \varphi_1(x_{it}) + \varepsilon_{1it} \quad (1-28)$$

$$y_{0it} = \varphi_0(x_{it}) + \varepsilon_{0it} \quad (1-29)$$

Où les termes d'erreurs sont supposés *iid*, c'est-à-dire $E(\varepsilon_{1it}|X_{it}) = 0$ et $E(\varepsilon_{0it}|X_{it}) = 0$. La variable x_{it} représente les caractéristiques observables et

indépendantes du traitement. On peut alors écrire le résultat observé de la manière suivante :

$$y_{1it} = \varphi_1(x_{it}) + D_{it}\alpha^* + \varepsilon_{1it} \quad (1-30)$$

Avec D_{it} la variable indicatrice indiquant la participation au programme et α^* représente l'effet du traitement, il est égal à la différence des deux résultats potentiels. Plusieurs méthodes sont utilisées pour estimer l'effet du traitement.

1.2.3.1 La méthode *Before-After* ou la méthode avant –après

L'estimateur de l'effet de traitement sur les traités par la méthode *Before-After* (Todd, 2008) correspond à la différence des résultats du groupe de traités après le traitement avec les résultats de ce même groupe avant le traitement. La méthode s'appuie sur une hypothèse postulant que le biais de sélection est en moyenne nulle pour toutes les personnes qui participent au programme.

Pour simplifier la notation, on suppose que l'effet du traitement est constant d'où :

$$\varphi_1(x_{it}) = \varphi_0(x_{it}) + \alpha^*$$

Soit t' et t désignent respectivement les périodes avant et après le programme. L'estimation du paramètre α^* peut être obtenue à partir de la différence première des résultats observés après et avant le programme.

$$Y_t - Y_{t'} = \varphi_0(x_t) - \varphi_0(x_{t'}) + \alpha^* + \varepsilon_t - \varepsilon_{t'} \quad (1-31)$$

Le modèle est estimé par les moindres carrés ordinaires et l'estimateur $\hat{\alpha}$ de α^* est interprété comme étant l'effet du programme sur les traités.

La mise en œuvre de cette méthode nécessite des données longitudinales, surtout la disponibilité des données sur les participants, au moins sur deux dates (ou deux périodes) : celle des données avant le programme et celle des données après le programme.

1.2.3.2 La méthode de régression « *cross section estimator* »

La méthode de régression « *cross section estimator* » est utilisée lorsque les données sur les individus traités ne sont pas disponibles avant le programme ou après le programme. Dans ce cas, pour estimer l'effet causal, les données des individus non traités sont utilisées pour construire le contrefactuel pour les individus traités. Elle consiste à estimer un modèle de régression classique avec une variable binaire (D_{it}) dans les explicatives indiquant la participation ou non de l'individu à la politique, tout en contrôlant les observables. La méthode repose sur l'hypothèse d'indépendance du traitement par rapport aux résultats observés (c'est-à-dire l'hypothèse CIA) comme dans les méthodes *matching*.

$$Y_t = \varphi_0(x_t) + D_t \alpha^* + \varepsilon_t \quad (1-32)$$

Ce modèle est estimé par les moindres carrés ordinaires et l'estimateur $\hat{\alpha}$ de α^* est interprété comme étant l'effet du programme sur les traités.

1.2.3.3 La méthode *Difference-in-Difference* (DID)

La méthode DID, aussi appelée « expérience naturelle » consiste en une différence entre l'estimateur *Before-After* des individus traités avec les individus non traités. Cette différence correspond à l'effet moyen du traitement sur les traités.

Soit un indicateur D qui est égal à 1, si l'individu i est traité et à 0, dans le cas contraire.

L'estimateur DID est un estimateur des moindres carrés ordinaires de α^* dans l'équation (1-33) :

$$Y_t - Y_{t'} = \varphi_0(x_t) - \varphi_0(x_{t'}) + D \alpha^* + \{\varepsilon_t - \varepsilon_{t'}\} \quad (1-33)$$

L'estimateur DID aborde la principale faiblesse de l'estimateur *Before-After* dans le sens où il permet d'introduire des paramètres d'effets temporels communs aux deux groupes (inclus dans $\varphi_0(x_{it})$) (Todd, 2008). La méthode DID suppose que la variation du

résultat sans le traitement est la même pour les traités que pour les non traités. Cette hypothèse signifie que les effets transitoires affectent de manière similaire à la fois les individus traités et les non traités. En plus des effets individuels et des effets transitoires qui caractérisent l'estimateur *Before-After*, il y a aussi les effets de tendance (effets temporels) qui varient dans le temps et sont communs aux individus traités et non traités. La mise en œuvre de la méthode DID nécessite des données transversales répétées (longitudinales), à la fois pour les traités et pour les non-traités.

L'avantage de la méthode DID est qu'elle permet d'éliminer les effets fixes individuels et les effets de tendance (comme dans l'estimateur *Before-After*) et par conséquent, seul l'effet transitoire peut influencer la décision de participer au programme, violant ainsi l'hypothèse d'identification.

C'est le cas s'il y a des comportements d'anticipation de certains traités par rapport au gain espéré ou lorsque la participation au programme est influencée par des chocs transitoires sur les résultats passés. Par exemple la participation à un programme de formation intervient suite à une baisse des revenus causée par une perte d'emploi ("*Ashenfelter's Dip*").

En conclusion, les différences entre les méthodes de régression et celles de l'appariement portent sur les hypothèses d'identification et la taille des échantillons (Jalan et Ravallion, 2003). Les méthodes de régression imposent des hypothèses que l'on peut qualifier d'arbitraires sur la forme fonctionnelle de l'effet du traitement (homogénéité de l'effet du traitement) et sur les variables de contrôle (linéarité), contrairement au PSM, où il y n' a aucune contrainte.

Par rapport à la taille de l'échantillon, les méthodes de régression utilisent l'ensemble des données disponibles alors qu'avec le PSM, seuls les individus de contrôle appariés sont conservés, les individus non-traités non appariés sont exclus de l'analyse. Comme le souligne Rubin and Thomas (2000), cité par Ravallion (2005, p.32), l'utilisation de l'ensemble des individus de contrôle implique que « *les estimations d'impact basées sur des échantillons complets (non appariées) sont généralement plus biaisées, et moins robustes que celles faites à partir d' échantillons appariés, à cause de la mauvaise spécification de la fonction de sélection* ».

Dans les sections 1.2.2 et 1.2.3 nous avons présenté les méthodes d'estimation de l'effet de traitement lorsque la sélection est basée sur les observables et nous avons posé les conditions dans lesquelles elles sont valides. Cependant, la sélection des individus ou encore la répartition de la population en groupe de traités et de non traités peut être également due aux caractéristiques inobservables. Nous exposons dans la section suivante, la méthode permettant d'estimer les paramètres d'intérêt dans ce cas.

1.2.4 Méthode d'estimation à partir de la sélection sur les inobservables

Lorsque le biais de sélection est dû aux inobservables, les méthodes d'estimation par *matching* et de régression sont moins fiables, car l'hypothèse *CIA* n'est plus valide et l'indépendance du traitement et de la participation n'est plus assurée. En effet, le biais de sélection est présent lorsqu'il existe des variables inobservables corrélées avec le traitement mais aussi avec la variable de résultat. C'est une forme d'endogénéité due à des variables omises.

A titre d'exemple, dans le cadre de l'évaluation d'une politique de formation des demandeurs d'emploi, l'économètre ne contrôle que les caractéristiques observables, mais ignore les *abilities* (aptitudes, compétences, capacités, talent, motivation...) du bénéficiaire (Goldberger, 2008). Si la formation offerte est corrélée avec ses *abilities* et que ces dernières sont corrélées avec les gains attendus, alors l'effet estimé est biaisé à cause de cette endogénéité. Par conséquent, l'omission des *abilities* du bénéficiaire entraîne une corrélation entre la variable de traitement (formation) et les résidus.

Pour contrôler le biais de sélection dû aux inobservables, Heckman (1979) propose une méthode d'estimation en deux étapes. Elle consiste à estimer, d'abord, une équation de sélection à travers un modèle de choix discret (probit, logit) afin de calculer, à partir des estimateurs obtenus, un terme de correction du biais appelé Inverse de Ratio de Mills (IMR). Dans la seconde étape, l'IMR est ajouté dans les équations de résultats comme une nouvelle variable explicative supplémentaire.

Cette approche, basée sur les données observables, permet d'estimer directement l'effet de traitement moyen sur la population ATE, à partir duquel on déduit les autres paramètres d'intérêt (ATT, ATU).

1.2.4.1. Spécification du modèle de sélection

Soit D_i^* une variable latente qui représente la propension d'un individu à être traité. Cette propension est fonction d'un ensemble de déterminants regroupés dans un vecteur de variables explicatives Z . Sous forme matricielle, le modèle structurel de la variable latente s'écrit de la manière suivante :

$$D^* = Z\theta + \varepsilon_D \quad (1-34)$$

Où θ est un vecteur de paramètre à estimer et ε_D représente le résidu suivant une distribution normale.

La variable latente D^* étant non observée, on observe une variable binaire D liée à D^* à travers la relation suivante :

$$D = \begin{cases} 1 & \text{si } D^* > 0 \\ 0 & \text{si } D^* \leq 0 \end{cases} \quad (1-35)$$

L'estimation de cette équation de sélection permettra de calculer l'Inverse du Ratio de Mills (IRM) qui sera intégré dans les équations de résultats.

1.2.4.2 Spécification des équations de résultats

Chaque groupe est caractérisé par une équation de résultats spécifiant que le résultat observé pour un individu, selon son statut, est une fonction linéaire de ses caractéristiques.

Pour les individus traités, c'est-à-dire $D = 1$, l'équation de résultats est donnée sous forme matricielle par :

$$Y_1 = X\beta_1 + \varepsilon_1 \quad (1-36)$$

Pour les individus non traités, c'est-à-dire $D = 0$, l'équation de résultat est donnée sous forme matricielle par :

$$Y_0 = X\beta_0 + \varepsilon_0 \quad (1-37)$$

Avec ε_1 et ε_0 , respectivement, les termes d'erreurs des équations de résultats avec les propriétés usuelles (ε_1 et $\varepsilon_0 \sim \mathcal{N}(0, \sigma^2 I)$), X la matrice des variables explicatives et β_1 et β_0 les vecteurs des paramètres à estimer.

En utilisant les équations (1-34), (1-36) et (1-37), le résultat observé peut s'écrire sous la forme générale suivante :

$$\begin{aligned} Y &= DY_1 + (1 - D)Y_0 \\ &= D(X\beta_1 + \varepsilon_1) + (1 - D)(X\beta_0 + \varepsilon_0) \\ &= DX\beta_1 + (1 - D)X\beta_0 + [\varepsilon_0 + D(\varepsilon_1 - \varepsilon_0)] \\ &= DX\beta_1 + (1 - D)X\beta_0 + u \end{aligned} \quad (1-38)$$

Avec $u = \varepsilon_0 + D(\varepsilon_1 - \varepsilon_0)$, un terme d'erreur composite, indépendant des variables explicatives, mais corrélé avec le traitement D , ce qui conduit naturellement à un biais d'endogénéité.

Dans le cadre d'une estimation en deux étapes à la Heckman (Heckman, 1979), on suppose que la distribution jointe des termes $(\varepsilon_D, \varepsilon_1, \varepsilon_0)$ suit une loi normale de moyenne 0 et de matrice de variance-covariance :

$$\begin{pmatrix} \varepsilon_D \\ \varepsilon_1 \\ \varepsilon_0 \end{pmatrix} \propto \mathcal{N} \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & r_{1\varepsilon_D}\sigma_1 & r_{0\varepsilon_D}\sigma_0 \\ r_{11}\sigma_1 & \sigma_1^2 & r_{10}\sigma_1 \\ r_{00}\sigma_0 & r_{01}\sigma_0 & \sigma_0^2 \end{pmatrix} \right]$$

Ainsi, pour tout k , avec $k = \{0,1\}$, on a :

$$E(Y_k|D = k) = X\beta_k + E(\varepsilon_k|D = k) \quad (1-39)$$

A partir de l'équation (1-39), pour tout k ,

$$E(\varepsilon_k|D = k) \neq 0 \quad (1-40)$$

Sauf si la corrélation $(\varepsilon_k, \varepsilon_D)$ est nulle c'est-à-dire $r_{k\varepsilon_D} = 0$.

En remplaçant k par ses valeurs dans l'équation (1-39), on obtient finalement.

Pour $k = 1$

$$\begin{aligned} E(\varepsilon_1|D = 1) &= r_{1\varepsilon_D}\sigma_1 E[\varepsilon_D|\varepsilon_D > -Z\theta] \\ E(\varepsilon_0|D = 1) &= r_{0\varepsilon_D}\sigma_0 E[\varepsilon_D|\varepsilon_D > -Z\theta] \end{aligned} \quad (1-41)$$

Pour $k = 0$

$$\begin{aligned} E(\varepsilon_1|D = 0) &= r_{1\varepsilon_D}\sigma_1 E[\varepsilon_D|\varepsilon_D \leq -Z\theta] \\ E(\varepsilon_0|D = 0) &= r_{0\varepsilon_D}\sigma_0 E[\varepsilon_D|\varepsilon_D \leq -Z\theta] \end{aligned} \quad (1-42)$$

Le terme final dans chaque expression est l'espérance d'une variable aléatoire normale tronquée où :

$$E[\varepsilon_D|\varepsilon_D > -Z\theta] = \frac{\phi(-Z\theta)}{1 - \Phi(-Z\theta)} = \frac{\phi(Z\theta)}{\Phi(Z\theta)}; \text{ pour } k = 1$$

$$E[\varepsilon_D|\varepsilon_D \leq -Z\theta] = -\frac{\phi(-Z\theta)}{\Phi(-Z\theta)} = -\frac{\phi(Z\theta)}{1 - \Phi(Z\theta)}; \text{ pour } k = 0$$

Les termes $\pi_1 = \frac{\phi(Z\theta)}{\Phi(Z\theta)}$ et $\pi_0 = \frac{\phi(Z\theta)}{1-\Phi(Z\theta)}$, représentent l'Inverse du Ratio de Mills servant à contrôler le biais de sélection dû aux inobservables et sont obtenus à partir de l'équation de sélection (1-34) et sont intégrés à l'équation de résultats.

Ainsi, l'espérance du résultat observé pour le groupe des traités est :

$$E(Y_1|D = 1) = X\beta_1 + E(\varepsilon_1|D = 1) = X\beta_1 + r_{1\varepsilon_D}\sigma_1 \frac{\phi(Z\theta)}{\Phi(Z\theta)} \quad (1-43)$$

L'espérance du résultat observé pour le groupe de traités s'il n'avait reçu le traitement est égale à :

$$E(Y_0|D = 1) = X\beta_0 + E(\varepsilon_0|D = 1) = X\beta_0 + r_{0\varepsilon_D}\sigma_0 \frac{\phi(Z\theta)}{\Phi(Z\theta)} \quad (1-44)$$

L'espérance du résultat observé pour le groupe de contrôle est :

$$E(Y_0|D = 0) = X\beta_0 + E(\varepsilon_0|D = 0) = X\beta_0 - r_{0\varepsilon_D}\sigma_0 \frac{\phi(Z\theta)}{1 - \Phi(Z\theta)} \quad (1-45)$$

L'espérance du résultat observé pour le groupe de contrôle s'ils n'avaient pas reçu le traitement est égale à :

$$E(Y_1|D = 0) = X\beta_1 + E(\varepsilon_1|D = 0) = X\beta_1 - r_{1\varepsilon_D}\sigma_1 \frac{\phi(Z\theta)}{1 - \Phi(Z\theta)} \quad (1-46)$$

1.2.4.3 Identification des paramètres d'intérêt

En remplaçant les équations (1-43), (1-44) et (1-45) dans les équations (1-4), (1-5) et (1-6), on obtient de nouveaux paramètres d'intérêt tenant compte du biais de sélection lié aux inobservables (Schroeder, 2010, p.216).

- **L'effet de traitement moyen sur les traités est donné par:**

$$\begin{aligned}
 ATT(X, Z) &= E(Y_1|D = 1) - E(Y_0|D = 1) \\
 &= X(\beta_1 - \beta_0) + [r_{1\varepsilon_D}\sigma_1 - r_{0\varepsilon_D}\sigma_0] \frac{\phi(Z\theta)}{\Phi(Z\theta)} \\
 &= X(\beta_1 - \beta_0) + \underbrace{[r_{1\varepsilon_D}\sigma_1 - r_{0\varepsilon_D}\sigma_0]}_{\text{biais de sélection dû aux inobservables}} \pi_1
 \end{aligned} \tag{1-47}$$

- **L'effet de traitement moyen sur les non traités est donné par :**

$$\begin{aligned}
 ATU(X, Z) &= E(Y_0|D = 0) - E(Y_1|D = 0) \\
 &= X(\beta_1 - \beta_0) - [r_{1\varepsilon_D}\sigma_1 - r_{0\varepsilon_D}\sigma_0] \frac{\phi(Z\theta)}{1 - \Phi(Z\theta)} \\
 &= X(\beta_1 - \beta_0) - \underbrace{[r_{1\varepsilon_D}\sigma_1 - r_{0\varepsilon_D}\sigma_0]}_{\text{biais de sélection dû aux inobservables}} \pi_0
 \end{aligned} \tag{1-48}$$

- **L'effet de traitement moyen sur la population est égal à :**

$$\begin{aligned}
 ATE(X, Z) &= P_r(D = 1|X, Z)ATT(X, Z) + P_r(D = 0|X, Z)ATU(X, Z) \\
 &= \Phi(Z\theta)ATT(X, Z) + 1 - \Phi(Z\theta)ATU(X, Z) \\
 &= X(\beta_1 - \beta_0) + [r_{1\varepsilon_D}\sigma_1 - r_{0\varepsilon_D}\sigma_0]\phi(Z\theta) - [r_{1\varepsilon_D}\sigma_1 - r_{0\varepsilon_D}\sigma_0]\phi(Z\theta) \\
 &= X(\beta_1 - \beta_0)
 \end{aligned} \tag{1-49}$$

A partir de cette dernière équation (1-49), on peut en déduire les paramètres d'intérêt (équations (1-47) et (1-48)) en fonction de ATE.

$$ATT = ATE + \pi_1[r_{1\varepsilon_D}\sigma_1 - r_{0\varepsilon_D}\sigma_0] \tag{1-50}$$

$$ATU = ATE + \pi_0[r_{1\varepsilon_D}\sigma_1 - r_{0\varepsilon_D}\sigma_0] \tag{1-51}$$

Les termes $\pi_1[r_{1\varepsilon_D}\sigma_1 - r_{0\varepsilon_D}\sigma_0]$ et $\pi_0[r_{1\varepsilon_D}\sigma_1 - r_{0\varepsilon_D}\sigma_0]$ représentent le biais de sélection dû aux inobservables, respectivement, dans les équations de résultats du groupe des traités (équation (1–50) et des non traités (1–51).

Le biais de sélection comprend deux éléments : l'un provenant de l'équation de résultat du groupe des traités et l'autre provenant de l'équation de résultat du groupe de contrôle. L'intensité de chaque élément augmente avec la corrélation entre les inobservables des équations de sélection et de résultats. Lorsque cette covariance est nulle, cela implique que les inobservables sont non corrélés et il y a absence du biais de sélection.

Cette méthode s'appuie, cependant, sur un ensemble d'hypothèses sur les équations de sélection et de résultats potentiels. L'Inverse du Ratio de Mills est valable seulement sous trois conditions : (i) l'équation de sélection est un modèle probit (ii) l'équation de résultat est linéaire et (iii) les termes d'erreurs des deux équations suivent une distribution normale bivariée. Lorsque ces conditions ne sont pas satisfaites, alors l'IMR correspondant n'est plus adapté pour contrôler le biais de sélection.

1.2.5 Méthode d'estimation par le « *bound nonparametric methods* »

La méthode d'estimation non-paramétrique par les bornes – que nous énoncerons par la suite sous son appellation usuelle *nonparametric bounds*, est introduite pour la première fois par Manski (1989 ; 1990). C'est une méthode qui présente, à l'instar des approches non-paramétriques, l'avantage de ne pas imposer d'hypothèses fortes pour l'estimation des paramètres d'intérêt ; en l'occurrence, les effets de traitement en évaluation des politiques publiques. Elle est adaptée lorsqu'on ne dispose pas suffisamment d'information sur le processus de sélection des individus et les auteurs qui ont utilisé cette méthode dans leurs études avancent essentiellement cette justification. On rappelle le cadre général. La définition de l'effet moyen de traitement total ne change pas. Il s'écrit comme dans le cadre du modèle de Rubin, c'est-à-dire :

$$\text{ATE} = E(Y_1|X) - E(Y_0|X) \quad (1-52)$$

En utilisant le théorème d'espérance *itérée* (The law of iterated expectations), on obtient, pour les individus traités et non traités :

$$E(Y_1|X) = E(Y_1|X, D = 1)P(D = 1|X) + E(Y_1|X, D = 0)P(D = 0|X) \quad (1-53)$$

$$E(Y_0|X) = E(Y_0|X, D = 0)P(D = 0|X) + E(Y_0|X, D = 1)P(D = 1|X) \quad (1-54)$$

Les termes $E(Y_1|X, D = 1)$, $E(Y_0|X, D = 0)$, $P(D = 1|X)$ et $P(D = 0|X)$ sont identifiables à partir des données.

En revanche les termes $E(Y_1|X, D = 0)$ dans l'équation (1-53) et $E(Y_0|X, D = 1)$ dans l'équation (1-54) sont inobservés.

L'idée principale de cette approche est que le résultat observé d'un individu est borné. Pour le groupe traité, Y_1 est compris dans l'intervalle Y_1^l et Y_1^u , représentant respectivement la borne inférieure et la borne supérieure du résultat observé. C'est également le cas pour le groupe de contrôle où la variable de résultat Y_0 est comprise dans l'intervalle Y_0^l et Y_0^u , représentant respectivement la borne inférieure et la borne supérieure du résultat.

Dans l'équation (1-53), on remplace donc la partie inobservée $E(Y_1|X, D = 0)$ par les bornes supérieure et inférieure de la variable résultat et on obtient :

$$\begin{aligned} E(Y_1|X, D = 1)P(D = 1|X) + Y_1^l P(D = 0|X) \\ \leq E(Y_1|X) \leq \\ E(Y_1|X, D = 1)P(D = 1|X) + Y_1^u P(D = 0|X) \end{aligned} \quad (1-55)$$

L'équation (1-55) implique que l'espérance conditionnelle du résultat $E(Y_1|X)$ est bornée par B_1^l et B_1^u représentant respectivement la borne inférieure et supérieure de $E(Y_1|X)$, avec :

$$\begin{aligned} B_1^l &= E(Y_1|X, D = 1)P(D = 1|X) + Y_1^l P(D = 0|X) \\ &= E(Y_1|X, D = 1)P(D = 1|X) + Y_1^l (1 - P(D = 1|X)) \end{aligned}$$

$$\begin{aligned} B_1^u &= E(Y_1|X, D = 1)P(D = 1|X) + Y_1^u P(D = 0|X) \\ &= E(Y_1|X, D = 1)P(D = 1|X) + Y_1^u (1 - P(D = 1|X)) \end{aligned}$$

Dans l'équation (1-54), on remplace la partie inobservée $E(Y_0|X, D = 1)$ par les bornes supérieure et inférieure de la variable résultat. On obtient :

$$\begin{aligned} E(Y_0|X, D = 0)P(D = 0|X) + Y_0^l P(D = 1|X) \\ \leq E(Y_0|X) \leq \\ E(Y_0|X, D = 0)P(D = 0|X) + Y_0^u P(D = 1|X) \end{aligned} \tag{1-56}$$

L'équation (1-56) implique que l'espérance conditionnelle du résultat $E(Y_0|X)$ est bornée par B_0^l et B_0^u représentant respectivement la borne inférieure et supérieure de $E(Y_0|X)$, avec :

$$\begin{aligned} B_0^l &= E(Y_0|X, D = 0)P(D = 0|X) + Y_0^l P(D = 1|X) \\ &= E(Y_0|X, D = 0)(1 - P(D = 1|X)) + Y_0^l P(D = 1|X) \end{aligned}$$

$$\begin{aligned} B_0^u &= E(Y_0|X, D = 0)P(D = 0|X) + Y_0^u P(D = 1|X) \\ &= E(Y_0|X, D = 0)(1 - P(D = 1|X)) + Y_0^u P(D = 1|X) \end{aligned}$$

Toutes choses égales par ailleurs, les équations (1-55) et (1-56), impliquent alors que l'effet moyen de traitement ATE est également borné par la borne inférieure par B_{ATE}^l et par la borne supérieure B_{ATE}^u et il s'écrit de la manière suivante :

$$[B_1^l - B_0^u] \leq ATE \leq [B_1^u - B_0^l] \quad (1-57)$$

$$B_{ATE}^l \leq ATE \leq B_{ATE}^u$$

La borne inférieure B_{ATE}^l est donnée par :

$$\begin{aligned} B_{ATE}^l &= E(Y_1|X, D = 1)P(D = 1|X) + Y_1^l (1 - P(D = 1|X)) \\ &\quad - E(Y_0|X, D = 0)(1 - P(D = 1|X)) - Y_0^u P(D = 1|X) \\ &= \{E(Y_1|X, D = 1) - Y_0^u\}P(D = 1|X) + \{Y_1^l - E(Y_0|X, D = 0)\}(1 - P(D = 1|X)) \end{aligned}$$

La borne supérieure B_{ATE}^u est donnée par :

$$\begin{aligned} B_{ATE}^u &= E(Y_1|X, D = 1)P(D = 1|X) + Y_1^u (1 - P(D = 1|X)) \\ &\quad - E(Y_0|X, D = 0)(1 - P(D = 1|X)) + Y_0^l P(D = 1|X) \\ &= \{E(Y_1|X, D = 1) - Y_0^l\}P(D = 1|X) + \{Y_1^u - E(Y_0|X, D = 0)\}(1 - P(D = 1|X)) \end{aligned}$$

Il faut noter que cette formulation initiale peut conduire à une définition d'un intervalle de valeur très grand, pouvant en outre contenir la valeur nulle ce qui ne permettrait pas de déterminer le signe de l'effet de traitement. Dans ce cas l'impact de la politique est indéterminé. Cette situation est désignée sous l'appellation « *worst-case bounds* » ou « *non assumption bounds* » (Manski, 1990).

Pour contourner ces limites et obtenir une estimation plus précise des bornes de l'effet de traitement et d'en déduire le signe de l'effet, Manski et Pepper (2000) posent deux hypothèses supplémentaires.

Soit un individu $i \in I$ avec I l'ensemble de la population, ayant un traitement $z_i \in T$ et un résultat observé $y_i := y_i(z_i) \in Y$. Le résultat potentiel $y_i(t)$ avec $t \neq z_i$ est inobservable (Manski et Pepper, 2000).

La première hypothèse est l'hypothèse MTS (*Monotone Treatment Selection*) qui suppose que les résultats des individus traités sont légèrement supérieurs à ceux des individus non traités.

Soit T un ensemble ordonné des traitements. Pour chaque $t \in T$, on a :

$$z_1 \geq z_0 \implies E[y(t)|z = z_1] \geq E[y(t)|z = z_0]$$

Par exemple, dans l'étude portant sur l'effet des études supérieures sur les salaires, l'hypothèse MTS postulée par Manski et Pepper (2000, p. 1002) suppose que « *persons who realize higher years of schooling have weakly higher mean wage functions than do those who realize lower levels of schooling* ».

La seconde hypothèse est l'hypothèse MTR (*Monotone Treatment Response*), qui suppose que le résultat est une fonction monotone du traitement, c'est-à-dire qu'il varie de manière croissante ou décroissante selon la nature du traitement. Cette hypothèse permet de déterminer le signe du traitement.

Soit T un ensemble ordonné des traitements. Pour chaque individu $i \in I$ alors :

$$t_1 \geq t_0 \implies y_i(t_1) \geq y_i(t_0)$$

Toujours dans leur application sur l'effet des études supérieures sur les salaires, l'hypothèse MTR suppose que « *ceteris paribus, wage rises as a function of conjectured years of schooling* » (Manski et Pepper, 2000, p.1002).

Cependant, comme dans les méthodes paramétriques, l'estimation de l'effet causal par cette méthode ignore les effets d'interaction dans l'estimation des bornes. Or la considération de ces effets d'interaction, s'ils sont présents, constitue un élément qui agit

sur les résultats et dans l'estimation des bornes. Par conséquent cela peut conduire à des conclusions en termes de politiques publiques plus appropriées. Nous proposons, dans le chapitre 5, une extension de la méthode avec la prise en compte des effets d'interaction dans l'estimation des bornes de l'effet de traitement.

Dans l'ensemble des méthodes présentées, les effets des interactions entre les individus sont ignorés que ce soit dans le processus de sélection ou dans l'estimation des paramètres d'intérêts. Or en réalité, les politiques, quel que soit le cadre, expérimental ou non expérimental, peuvent engendrer des effets résultant des interactions entre les individus (bénéficiaires et/ou non bénéficiaires). Dans le processus de sélection des bénéficiaires dans le cadre d'une politique non expérimentale, la décision de participation d'un individu peut être influencée par la décision des autres individus qui lui sont proches géographiquement ou qui appartiennent au même groupe social... C'est le cas par exemple des politiques « antitabac » (décision d'arrêter de fumer), des politiques fiscales locales (adoption d'une nouvelle taxe par une commune)...

D'autre part, selon le cadre - expérimental ou non expérimental - le résultat observé dans un territoire, dans un endroit ou chez un individu peut être également impacté par les résultats observés dans le voisinage immédiat. De ce fait, nous pouvons valablement supposer que le traitement reçu par un individu peut effectivement affecter les résultats des autres individus. Une politique de formation des demandeurs d'emploi qui vise à favoriser leur retour en emploi va impacter à la fois la possibilité pour chaque bénéficiaire et non bénéficiaire de retrouver un emploi en augmentant leur chance de retour à l'emploi ou en créant plus de concurrence entre les demandeurs d'emploi. Dans le cas d'une politique de soutien scolaire qui cible une partie d'une classe ou école, il est implicitement attendu que la politique améliore non seulement le niveau des élèves bénéficiaires, mais aussi celui des autres élèves non bénéficiaires de la classe ou de l'école du fait d'une amélioration des conditions d'apprentissage de tous.

La prise en compte des interactions entre les individus a trois implications dans l'évaluation des effets des politiques.

Premièrement, elle nécessite de s'interroger sur la validité des hypothèses retenues dans les modèles d'évaluation. C'est notamment le cas pour l'hypothèse SUTVA (*Stable*

Unit Treatment Value Assumption) qui est remise en cause dans le cadre expérimental (Hudgens et Halloran, 2008 ; Manski, 2012 ; Rosenbaum, 2007 ; Sobel, 2006 ; Tchetgen et VanderWeele, 2012) et dans le cadre non expérimental (Ferracci, 2010 ; Imbens et Wooldridge, 2009 ; Tchetgen et VanderWeele, 2012).

Deuxièmement, lorsque les effets d'interaction existent et ne sont pas traités alors l'effet de traitement estimé par les méthodes des moindres carrés ordinaires est affecté par des problèmes économétriques. En effet, la présence d'interdépendances dans le processus de sélection et des interactions dans les résultats peuvent poser un problème d'endogénéité qui se décline sous trois formes : variables omises- simultanéité – erreurs de mesure (Soetevent, 2006). Les résultats des estimations sont alors biaisés.

Troisièmement, les résultats obtenus conduisent à des conclusions erronées sur l'effet réel de la politique puisque les évaluations ne tiennent pas compte des effets induits du fait des externalités associées aux interactions. Les décisions prises par les décideurs politiques au vu des résultats de la politique ne sont pas prises dans des conditions optimales.

Quelles sont plus précisément ces interactions, leurs caractéristiques et leurs effets ? Ces points font l'objet de la section suivante (1.3) dans laquelle nous nous interrogeons sur les formes d'interactions, leurs sources et les mécanismes qui affectent les effets des politiques publiques.

Comment peut-on modéliser les interactions dans les modèles d'évaluation des effets de traitements ? Quelles méthodes économétriques permettent d'estimer l'impact de ces interactions dans les effets de traitement ? Ce point est présenté dans la dernière section (1.4) du chapitre.

1.3 Effets d'interaction et évaluation de l'effet de traitement

L'analyse des interactions est un sujet complexe qui dépasse le cadre de cette thèse. Nous souhaitons ici souligner les éléments qui se rapportent plus directement au cadre des politiques publiques.

1.3.1. Les institutions en tant que contexte des politiques publiques et des interactions

Les politiques publiques sont naturellement caractérisées par l'environnement dans lequel elles sont mises en œuvre, c'est-à-dire le contexte institutionnel et géographique sous lesquels vivent les bénéficiaires. Cet environnement n'est pas neutre et peut affecter une politique tant sur le processus de sélection que sur sa performance en termes d'effets sur les bénéficiaires.

Le contexte institutionnel peut être associé à la présence d'un ensemble d'institutions au sens de North (1991), qui régissent et encadrent l'organisation économique, sociale et politique. Ces institutions sont définies comme étant un ensemble de « *contraintes établies par les hommes qui structurent les interactions sociales, économiques et politiques* » North (1991, p.97). Ces contraintes peuvent être formelles (lois, modes de gouvernance ...) ou informelles (les coutumes et traditions, les codes de conduites, les sanctions...).

Les institutions formelles sont structurées autour et par un ensemble de textes et lois qui régissent le fonctionnement économique, juridique et politique d'une société. Ces institutions sont plus précisément regroupées en deux catégories. Les institutions économiques déterminent les règles et les contraintes sur les acteurs économiques et façonnent les résultats économiques. Les institutions politiques jouent le même rôle dans la sphère politique (Acemoglu et *al.* 2005).

Les institutions informelles à *la North* sont constituées de normes, de règles et de codes de conduites qui s'appliquent à des individus ou à des groupes d'individus. A cet égard, les membres d'un groupe sont tenus de se conformer aux règles. Ils doivent se conformer aux ordres de la hiérarchie, adopter les modèles proposés par leurs pairs à travers l'imitation, l'apprentissage, ou la coercition. Les individus appartenant à un même groupe ont tendance à avoir des comportements similaires, les mêmes opinions et les mêmes croyances.

Dans le cadre d'une politique publique, les institutions ont un rôle important à différents égards : dans l'adoption de la politique par les individus, au niveau de sa performance et de ses effets à travers les différentes formes d'interactions qu'elles génèrent.

Les interactions peuvent ainsi être liées à des comportements stratégiques face aux conditions économiques politiques et sociales. Elles sont conditionnées par les relations entre les membres d'un groupe, par les règles qui y sont instituées et elles relèvent de l'influence de certains individus.

Les termes d'interactions sociales sont habituellement utilisés pour désigner ces interactions et sont définies d'une manière générale comme les interactions entre les comportements individuels et les caractéristiques d'un groupe (Manski, 1993 ; Topa, 2001). Elles se déclinent sous différentes appellations et formes telles que les effets de voisinage, les effets de pairs, les effets d'imitation... Elles peuvent d'un côté, amplifier la portée et la réussite d'une politique et de l'autre côté, conduire à son échec.

Les politiques publiques, en tant que produit des institutions peuvent également en retour modifier le contexte institutionnel et agir sur les interactions. A cet égard, Moffit (2001) catégorise les politiques publiques en fonction de leurs effets sur les interactions sociales en trois catégories.

- La première est relative aux politiques qui modifient la composition d'un groupe en introduisant d'autres individus. Par exemple, l'accueil des élèves des zones défavorisées dans les écoles situées dans des zones favorisées ou bien des politiques urbaines de mixité sociale consistant à loger des ménages défavorisés dans des quartiers « riches ».
- La deuxième catégorie comprend les politiques qui modifient les caractéristiques d'un sous-ensemble d'un groupe. Il s'agit des politiques qui vont modifier quelques caractéristiques des individus traités, mais vont engendrer des effets importants sur les autres personnes non traités. C'est le cas par exemple des politiques d'exonération fiscales, des politiques de subventions...
- Enfin, la troisième catégorie concerne les politiques visant à agir directement sur les normes sociales en favorisant de manière durable de nouveaux comportements comme les politiques de contraception des jeunes filles, les politiques de lutte contre le tabagisme ...

1.3.2 Les interactions : définition et principaux mécanismes

Il est généralement fait référence à la catégorisation proposée par Manski (1993) pour classer ces différentes formes d'interactions en trois groupes. Pour illustrer la distinction entre ces différents effets nous reprenons l'exemple présenté par Manski (1993) qui porte sur la réussite scolaire des collégiens.

- (i) *Les interactions endogènes ou effets endogènes* apparaissent lorsque le comportement d'un individu varie avec le comportement des autres membres de son groupe. La réussite d'un collégien varie avec la réussite des autres élèves de sa classe, de son groupe ethnique, ou d'un autre groupe de référence.
- (ii) *Les interactions contextuelles ou effets de contexte* se produisent lorsque le comportement d'un individu varie seulement avec les caractéristiques des autres membres de son groupe. La réussite scolaire du collégien varie avec, par exemple, la composition socio-économique (revenus, taille des familles, âge ...) du groupe de référence.
- (iii) *Les effets corrélés* se produisent lorsque les individus d'un groupe adoptent des comportements similaires, car ils ont des caractéristiques similaires ou lorsqu'ils sont soumis aux mêmes institutions ou à des institutions similaires. Dans ce cas, les collégiens dans la même école ont tendance à réussir de façon similaire parce qu'ils ont contextes familiaux similaires ou parce qu'ils sont formés par les mêmes enseignants.

L'importance de ces effets, dans l'appréciation des effets d'une politique publique, dans les enseignements qui pourront être utilisés ensuite et au final dans la crédibilité de l'action publique, est aujourd'hui largement admise. Comme le souligne Manski (2000, p. 24) « *distinguishing among endogenous interactions, contextual interactions, and correlated effects has been thought important because these hypotheses imply different predictions for the impact of public policy* ».

Shipan et Volden (2008) indiquent, dans le cadre des politiques non expérimentales, que ces interactions se produisent à travers quatre principaux mécanismes, dont l'étude permet de mettre en exergue les processus par lesquels les individus interagissent entre eux.

- Le premier est l'apprentissage ou *learning*, qui repose sur le principe selon lequel la probabilité d'un individu de participer à une politique augmente lorsque les autres individus de son voisinage participent à la même politique ou ont déjà participé à la même politique.
- Le deuxième mécanisme porte sur la *compétition ou concurrence* qui suppose que la probabilité d'un individu de participer à une politique diminue lorsqu'elle a eu des

retombées négatives sur les individus de son voisinage, tandis qu'elle augmente si elle a des retombées positives sur son voisinage. Il s'agit, dans ce cas, des réactions ou des anticipations de l'individu par rapport aux comportements de ses voisins, afin de protéger ou de promouvoir ses intérêts. C'est le cas par exemple des politiques fiscales, des politiques de santé publiques (tabac, alcoolisme...).

- Le troisième mécanisme est relatif à l'*imitation*, qui repose sur le principe selon lequel la probabilité d'un individu d'adopter une politique augmente lorsque son plus proche (ou plus influent) voisin participe à la même politique ou a déjà participé à une politique similaire. C'est le cas par exemple des politiques d'innovation technologique où il y a les leaders et les suiveurs.
- Enfin, la *coercition* est un mécanisme présent lorsque le choix effectué par un individu est imposé par d'autres acteurs puissants (très influents) : nous retrouvons le cas des leaders sur les suiveurs, mais aussi par exemple l'influence d'une grande métropole sur les petites communes aux alentours, d'une grande entreprise avec des concurrents de petites tailles,...

Le contexte des interactions sociales permet par ailleurs de spécifier la portée de ces interactions qui sont implicitement attachées à des partitions sociales : les individus d'un groupe interagissent entre eux. La notion de voisinage traduit à la fois la proximité (*neighborhood*) et l'existence de liens sociaux entre voisin (*neighboring*). Les interactions sociales s'expriment implicitement aux seins de groupes délimités et leurs effets se manifestent aux seins de ces groupes. On peut alors faire l'analogie avec les groupes des traités et des non traités pour les politiques publiques. Cependant, un cadre plus général peut être proposé. En mobilisant la littérature sur la ségrégation sociale et la mixité sociale, Galster (2007), à la suite de plusieurs de ses travaux, propose de rapprocher les mécanismes sociaux évoqués précédemment d'une typologie des interactions selon la portée des interactions. Partant d'une situation sociale séparant les individus Riches (R) et les individus Pauvres (P), il s'interroge notamment sur les interactions possibles aux seins et entre les groupes. Il convient alors de tenir compte de trois éléments.

- Premièrement, l'effet positif ou négatif des interactions. Il est ainsi généralement postulé ou admis que les interactions au sein et avec le groupe R produiraient des effets positifs, tandis que les interactions au sein et avec le groupe P seraient négatives. Les

neighborhood effects par exemple illustrent la situation selon laquelle des externalités négatives au sein de groupes défavorisés accentuent les difficultés rencontrées par les membres de ce groupe (échec scolaire, chômage ...).

- Deuxièmement, la symétrie ou non des interactions intergroupes. Il est généralement postulé que cet effet est asymétrique et qu'il concerne plutôt les interactions que le groupe P souhaiterait avoir avec le groupe R ; ces interactions étant positives.
- Troisièmement, les effets des interactions peuvent être liés à une composition sociale mixte du voisinage et dépendent dans ce cas de la proportion \mathcal{D} des individus P présents dans ce voisinage. Il est généralement admis que les interactions positives – issues des individus R - peuvent se produire pour l'ensemble du voisinage en dessous d'une proportion \mathcal{D}' tandis que les interactions négatives – issues des individus de type P, se manifestent dès et au-delà d'une proportion \mathcal{D}'' .

Le signe, le nombre et les types d'individus qui interagissent sont alors à considérer. Sont directement attachés à ces types d'interactions les effets de voisinage négatifs au sein des groupes défavorisés notamment, les effets de stigmatisation, les effets de normes sociales associé à l'existence d'un nombre suffisant d'individus pour que se développe la norme, les effets de compétition lorsque les effets négatifs sont exercés par un groupe vis-à-vis d'individus de l'autre groupe...

Les effets décrits dans ce cadre abordent les interactions entre groupes sociaux et autorisent la composition mixte du voisinage. C'est en précisant la notion de voisinage que sont abordées généralement les caractéristiques géographiques des interactions.

1.3.3. Le contexte géographique des interactions

Outre les interactions liées au contexte institutionnel, une politique publique peut également être confrontée aux interactions liées au contexte spatial ou géographique dans lequel elle est mise en œuvre.

Le contexte géographique fait référence principalement à la zone d'action de la politique. Il est alors caractérisé par le périmètre d'action – au sein d'une zone bien délimitée – et par le fait que les actions seront localisées en certains lieux relativement à d'autres lieux. En effet, les actions d'une politique publique sont soit mises en œuvre sur un territoire précis – on parle alors de politiques zonées ou territoriales – soit elles sont

destinées à des individus, qui par nature sont localisés dans des endroits bien définis. Ce contexte géographique peut être caractérisé par des variables géographiques (distance, périmètre, altitude, relief ...).

Dans l'évaluation des politiques publiques, ces caractéristiques sont prises en compte par exemple pour analyser l'impact de l'éloignement ou l'impact de l'enclavement sur les résultats de la politique. Ces caractéristiques géographiques ne sont pas neutres, cependant, quant aux interactions et à leurs effets. Les frictions aux interactions associées à ces variables géographiques peuvent limiter l'étendue des interactions et les amener à se concentrer dans des périmètres précis. A titre d'exemple, l'enclavement/désenclavement de la zone d'action peut impacter positivement la diffusion des effets d'une politique (diffusion d'une nouvelle technologie agricole, programme de vaccination). Pour Briant et al. (2014, p. 24) *« deux dimensions géographiques ont notamment conditionné le succès du programme ZFU en France. L'accessibilité des transports, et plus particulièrement l'accès à la route et le train ou les stations de métro, améliore la capacité du programme à attirer des entreprises et créer des emplois. En revanche, les discontinuités générées par la circulation ou des barrières naturelles entre le quartier et les principaux centres d'emploi de son aire urbaine entravent les avantages d'allègements fiscaux et des exonérations sociales »*.

Dans le cadre d'une politique d'emploi, l'éloignement du périmètre d'action des zones d'emplois ou mal desservies par les réseaux de transport en commun, va nécessairement impacter la décision des potentiels bénéficiaires et l'effet de la politique sur le retour en emploi des bénéficiaires (Bunel et al. 2015).

Par extension, le contexte géographique peut être assimilé à l'ensemble des caractéristiques structurelles (démographie, infrastructures, économiques) et politiques (fiscales, économiques, sociales) de la zone d'application affectant les comportements des individus. Il n'est donc pas neutre quant à la nature des interactions sociales qui peuvent s'y tisser et aux impacts de ces interactions.

Le contexte géographique renvoie finalement au concept de voisinage (son étendue, sa situation vis-à-vis des autres espaces et ses caractéristiques socio-économiques et institutionnelles). Le contexte spatial décrit plus généralement les

mécanismes spatiaux qui vont régir la localisation des individus, la concentration géographique, la diffusion spatiale ...

Ces deux aspects intéressent plus spécifiquement les effets des interactions sur les politiques publiques.

Selon Haining (2004), les interactions entre les individus dans les zones d'actions et en dehors, s'opèrent à travers quatre types de mécanismes spatiaux, par lesquels les individus interagissent mais aussi par lesquels se diffusent les effets d'une politique.

- Le processus de *diffusion spatiale* est présent lorsqu'un « attribut » est adopté par une partie de la population. Haining (2004, p. 21) précise que la propagation de l'attribut au sein de la population dépend de l'attribut lui-même et souligne que « *l'acquisition ou l'adoption consciente ou inconsciente de l'attribut peut dépendre de contacts interpersonnels, de la communication ou de l'influence et de la pression* ».
- Le second processus *échange et transfert* résulte du fait que les économies des unités spatiales sont liées entre elles. Par exemple, « *le revenu gagné dans la production et la vente d'un produit dans un endroit peut être dépensé en biens et services dans d'autres endroits* » Haining (2004, p.22).
- Le troisième mécanisme, à savoir le processus *d'interaction spatiale*, est présent lorsque les résultats observés dans une zone d'application du programme influencent - et sont influencés - par les résultats observés dans les autres zones hors programme. Par exemple la construction d'une école dans un village permet non seulement d'améliorer le taux de scolarisation dans ce village mais aussi celui des autres villages voisins. En revanche, les villages non bénéficiaires peuvent influencer négativement à leur tour le village bénéficiaire en les convainquant de l'inadéquation de ce programme par rapport à leurs valeurs sociales.
- Enfin, le dernier type de processus est le processus de *dispersion* qui est attaché à la dispersion de la population elle-même dans l'espace. Par exemple, l'adoption d'une nouvelle espèce agricole à haute rendement par un village peut se propager dans toute la région à travers divers vecteurs comme les marchés, les déplacements des personnes, les rencontres familiales ...

Cependant, ces effets d'interaction sont bien limités dans un voisinage défini. C'est souvent posé par hypothèse dans le cadre de politiques zonées et donc instauré dans

l'évaluation de la politique : les effets de la politique seront supposés nuls en dehors du voisinage. Sans faire référence à un périmètre bien délimité, la portée spatiale des interactions est souvent délimitée par la distance au-delà de laquelle on suppose qu'il n'y a plus d'interactions entre les individus. Par conséquent l'effet de la diffusion de la politique est supposé nul au-delà de cette distance. Bien sûr, les évaluations peuvent reconsidérer ces hypothèses et le font généralement en tenant compte des effets de frontières ou de portées plus ou moins grandes. Il s'agit alors plus de caractériser l'étendue des interactions que d'intégrer précisément les sources des interactions.

A cet égard, le concept de voisinage renvoie à la notion de proximité. Dans un contexte social, le voisinage est à la fois le lieu où habitent les voisins et l'endroit où sont confinées les interactions. La proximité renseigne à la fois sur la sphère d'influence des individus les uns par rapport aux autres, sur l'ampleur et les mécanismes de leurs interactions. Il est admis que la proximité va favoriser les interactions. Elle peut être définie de deux manières.

Le plus communément, la proximité désigne la proximité géographique qui est mesurée par la distance entre deux individus, exprimée en distance métrique ou par toute autre forme corrélée à la longueur du déplacement comme le temps de transport. Le voisinage géographique entendu comme un petit périmètre défini généralement par des contiguïtés rentre aussi dans cette catégorie. Une mesure de la friction spatiale, c'est-à-dire des freins aux interactions qui seraient associés à de l'éloignement géographique sont donc considérés à travers la proximité géographique.

Le second type de proximité est associé à d'autres types de frictions aux échanges et fait référence à tout ce qui peut favoriser ou freiner les interactions. Ça peut-être des critères économiques (proximité économique), critères institutionnels (proximité institutionnelle), critères culturels (proximité socio-culturelle).

- La proximité économique fait référence à la situation dans laquelle des individus ou unités spatiales ont des caractéristiques économiques similaires (Baumont et Guillain, 2013). Pour certains types de politiques, la proximité économique peut mieux appréhender le voisinage que la proximité géographique. C'est le cas de certaines politiques fiscales où la proximité peut être définie par la similarité du PIB par habitant,

pour les politiques commerciale (les flux commerciaux bilatéraux), certaines politiques territorialisées (les dépenses publiques en R&D).

- La proximité institutionnelle indique que les unités spatiales ont des institutions similaires (Baumont et Guillain, 2013 ; Plaigin, 2012). Elle permet dans certains cas de mieux spécifier les interactions entre les individus. Elle lie les individus soumis aux mêmes institutions telles que l'appartenance à une même région, à un même EPCI (Etablissement Public de Coopération Intercommunale). Par exemple, l'appartenance à un même EPCI incite les communes membres à adopter le même comportement vis-à-vis des actions communautaires.

- La proximité culturelle/sociale (Leenders, 2002 ; Plaigin, 2012) indique que les unités spatiales ont des caractéristiques sociales (structure par âge, CSP, éducation, santé ...) et/ou culturelles similaires (même langue, même religion, même groupe ethnique...). La proximité culturelle permet de traduire convenablement les interactions entre les individus dans certains types de politiques. Par exemple les politiques visant à modifier certaines valeurs sociales de certaines communautés, afin de pouvoir atteindre leurs objectifs, vont utiliser l'importance des relations sociales pour apprécier les proximités.

Face à l'importance que la prise en compte des interactions peut apporter à une appréciation plus réaliste des mécanismes à l'œuvre dans les groupes ou dans les territoires, il semble important de proposer un cadre d'évaluation des politiques publiques qui les intègre. Par ailleurs, les effets des interactions viennent s'ajouter aux effets produits par la politique. Il convient alors de proposer un cadre d'évaluation qui tienne compte également des effets de voisinage. Pour construire un tel cadre, tant sur le plan économétrique qu'économique, un certain nombre de questions se pose. En quoi la prise des interactions modifie le cadre conceptuel de l'évaluation des effets de traitement ? Comment prendre en compte les effets d'interaction dans la définition de l'effet de traitement ? Quels types d'interactions considérer ? Au sein des groupes ? Entre les groupes ? Quels types de voisinage ? Enfin, à quels niveaux les interactions doivent être considérées dans l'évaluation de la politique : au niveau du processus de sélection ? Au niveau des résultats ? A chacune de ces étapes ?

Pour répondre à ces questions, nous proposons une approche basée sur les méthodes de l'économétrie spatiale pour intégrer les effets d'interaction dans l'estimation de l'effet de la politique. Avant de présenter dans la section 1.4 les modèles économétriques que nous utiliserons dans la suite de la thèse, le prochain paragraphe rappelle quelques fondements de la prise en compte des interactions dans l'évaluation des politiques publiques.

1.3.4 Les interactions dans l'évaluation des effets des politiques publiques

1.3.4.1 La modélisation des interactions

L'économétrie spatiale offre un cadre approprié pour spécifier les voisinages et modéliser les interactions spatiales. Sur ce dernier point, la classification des interactions sociales proposée par Manski (1993) peut être formalisée par des spécifications spatiales (Soetevent, 2006).

- Les effets endogènes sont modélisés à travers une spécification de type *spatial autoregressive model* (SAR).
- Les effets exogènes sont estimés à travers une spécification *Spatial Durbin model* (SDM)
- Les effets corrélés sont estimés à l'aide d'un modèle *spatial error model* SEM.

Le Tableau 1-1 synthétise ces formes d'interactions, simples ou combinées, et leur modélisation via les spécifications spatiales. Les spécifications peuvent être appliquées à l'équation de sélection et à l'équation de résultat. Les effets des interactions sont alors estimés via les méthodes de l'économétrie spatiale.

Les matrices W de pondérations spatiales permettent de spécifier les schémas d'interdépendances supposés relier les individus. Ces schémas définissent à la fois l'intensité des connexions entre chaque paire d'individus et l'étendue du voisinage sur lesquelles ces connexions se produisent. Les termes w_{ij} de la matrice de pondération traduisent donc d'une part, si les individus i et j sont dans le même voisinage ($w_{ij} \neq 0$) ou pas ($w_{ij} = 0$) et d'autre part, si les individus i et j interagissent plus ou moins en

fonction de leur degré de proximité : les poids w_{ij} sont définis par une fonction non décroissante du degré de proximité entre i et j . Par exemple, si la distance d_{ij} est supérieure à la distance d_{ik} , alors l'individu i interagit plus avec l'individu k qu'avec l'individu j : $w_{ik} > w_{ij}$.

L'estimation des paramètres spatiaux - ρ pour un effet spatial autorégressif, λ pour l'autocorrélation des erreurs ou θ pour un effet spatial sur les variables exogènes – fournit les éléments nécessaires à l'estimation de l'impact des interactions sur l'effet de traitement et donc à l'évaluation de cet effet de traitement.

Si les interactions ne sont pas prises en compte alors qu'elles sont présentes, l'estimation moindres carrés ordinaire des paramètres ne sont pas fiables : les estimateurs sont inefficients en présence d'autocorrélation spatiale des erreurs et ils sont biaisés et non convergents dans le cas autorégressifs (LeSage et Pace, 2009).

Tableau 1-1: Différents modèles pour l'estimation de la performance de la politique selon les types d'effets

Modèle	hypothèses	Modèle	hypothèses	Modèle	hypothèses	Modèle
GNS General Nesting Spatial $Y = \rho WY + X\beta + XW\theta + u$ $u = \lambda Wu + \varepsilon$	$\theta = 0$	SAC $Y = \rho WY + X\beta + u$ $u = \lambda Wu + \varepsilon$	$\lambda = 0$	SAR $Y = \rho WY + X\beta + u$	$\rho = 0$	OLS $Y = X\beta + u$
			$\rho = 0$	SEM $Y = X\beta + u$ $u = \lambda Wu + \varepsilon$	$\lambda = 0$	
	$\lambda = 0$	SDM $Y = \rho WY + X\beta + XW\theta + u$	$\theta = 0$	SAR $Y = \rho WY + X\beta + u$	$\rho = 0$	
			$\rho = 0$	SLX $Y = X\beta + XW\theta + u$	$\theta = 0$	
			$\theta = -\rho\beta$	SEM $Y = X\beta + u$ $u = \lambda Wu + \varepsilon$	$\lambda = 0$	
	$\rho = 0$	SDEM $Y = X\beta + XW\theta + u$ $u = \lambda Wu + \varepsilon$	$\lambda = 0$	SLX $Y = X\beta + XW\theta + u$	$\theta = 0$	
			$\theta = 0$	SEM $Y = X\beta + u$ $u = \lambda Wu + \varepsilon$	$\lambda = 0$	

Source : Adapté de Elhorst (2010, p. 13)

1. Les modèles SAR représentent des modèles à effets endogènes
2. Les modèles SEM représentent des modèles à effets corrélés
3. Les modèles SDM représentent des modèles à effets de contextes
4. Les modèles SAC représentent des modèles à effets endogènes et à effets corrélés
5. Les modèles SDEM représentent des modèles à effets de contextes et à effets corrélés

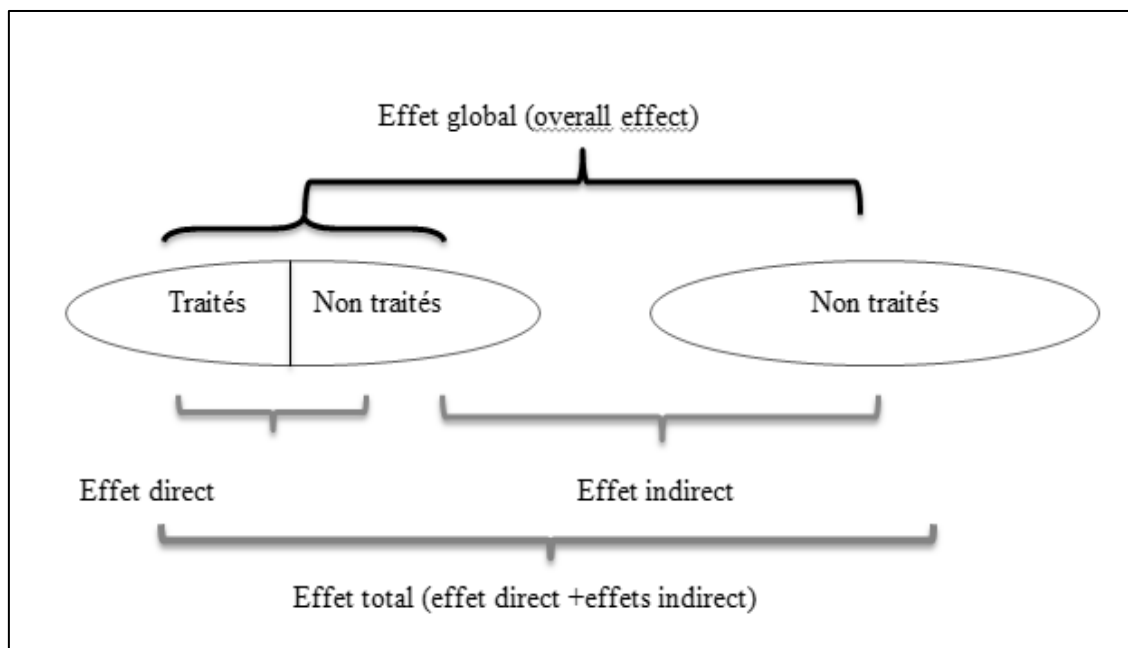
1.3.4.2 Implications sur les hypothèses des modèles d'évaluation

L'hypothèse SUTVA est postulée pour signifier l'absence d'interdépendance entre les bénéficiaires. Depuis quelques années, des approches statistiques (Hudgens et Halloran, 2008 ; Sobel, 2006) sont proposées pour lever partiellement l'hypothèse SUTVA et par conséquent prendre en compte, en partie les effets d'interaction dans l'estimation de l'effet de traitement. Ces méthodes s'appuient sur l'hypothèse d'interférence partielle - *Partial Interference Assumption* (PIA) ou *Control Treatment Response* (CTR). Elle postule que le traitement reçu par un individu appartenant à un groupe, affecte le résultat de cet individu mais aussi le résultat des autres membres du groupe **mais** il n'affecte pas le résultat des individus en dehors de son groupe.

Halloran et Struchiner (1991) et Hudgens et Halloran (2008) définissent deux groupes d'individus. Un premier groupe composé d'individus traités et des individus non traités et un deuxième groupe composé uniquement d'individus non traités (*cf.* Figure 1-2 :). Le premier groupe correspond, par exemple, à la population suite au processus de désignation, tandis que le second groupe correspond au groupe de contrôle pour l'évaluation du résultat.

L'effet de traitement peut alors être décomposé en quatre types d'effets : un effet direct, un effet indirect, un effet total et un effet global. L'effet direct est défini comme étant la différence entre le résultat potentiel d'un individu traité et le résultat potentiel d'un individu non traité appartenant au premier groupe. L'effet indirect est défini comme étant la différence du résultat potentiel d'un individu non traité appartenant au premier groupe et le résultat potentiel d'un individu non traité appartenant au deuxième groupe. L'effet total est défini comme étant la différence du résultat potentiel d'un individu traité et le résultat potentiel d'un individu non traité appartenant au deuxième groupe. En d'autres termes, l'effet total est la somme de l'effet direct et de l'effet indirect. L'effet global représente la moyenne de l'effet du traitement par rapport à une absence de traitement.

Figure 1-2 : La décomposition de l'effet de causal (Halloran et Struchiner, 1991)



Source : adapté de (Halloran et Struchiner, 1991)

En plus de l'hypothèse CTR, Manski (2013) propose une hypothèse supplémentaire sur les implications des interactions entre les individus d'un même groupe. Cette hypothèse appelée SMTR (*Semi-Monotone Treatment Response*) suppose que le résultat individuel varie avec le traitement qu'il reçoit et le traitement des autres membres de son groupe. L'hypothèse permet de définir empiriquement deux types d'effets d'interaction, à savoir les interactions de renforcement « *Reinforcing interaction* » et les interactions antagonistes « *Opposing interaction* ». Le cas « *reinforcing interaction* » signifie que le résultat d'un individu augmente à la fois avec le traitement qu'il reçoit et les traitements des autres membres de son groupe (c'est par exemple le cas de la vaccination). Le cas « *Opposing interaction* » signifie au contraire que le résultat d'un individu augmente avec le traitement qu'il reçoit, mais diminue avec les traitements des autres membres de son groupe (c'est le cas par exemple de la formation des demandeurs d'emploi sur le marché du travail).

Cependant, ces méthodes alternatives présentent deux limites majeures : (i) elles sont plus appropriées dans le cadre expérimental (Tchetgen et VanderWeele, 2012) et (ii)

elles s'appuient sur une hypothèse moins restrictive que SUTVA, mais limitant les interactions uniquement à l'intérieur du groupe.

Elles ne permettent pas de prendre en compte les interactions inter groupes, c'est-à-dire les interactions entre deux individus de groupes différents (un individu traité avec des individus non traités et vice-versa). Dans le cadre d'une politique publique non expérimentale, un individu bénéficiaire du traitement peut être en relation, au sein de son voisinage traité (groupe traité) ou en dehors de son voisinage non traité (groupe non traités), avec les deux types d'individus : des individus qui comme lui reçoivent le traitement et des individus non bénéficiaires et avec qui il entretient des relations de proximité. Il semble, en effet, restrictif de supposer que les individus n'interagissent qu'entre semblables. Par ailleurs, les interactions entre groupes différents relèvent aussi des interactions de voisinages comme nous l'avons présenté auparavant. Enfin, l'étendue du voisinage comme les formes de proximité ne conduisent pas à se restreindre sur seulement les interactions intra groupes. Cela supposerait un contrôle parfait des échanges ce qui est peu réaliste. Ignorer les potentielles interactions intergroupes, c'est-à-dire des interactions entre des individus traités avec des individus non traités, peut conduire à des estimations biaisées de l'effet de traitement.

Le cadre formel que nous proposons pour intégrer les effets des interactions dans l'évaluation de l'effet de traitement est maintenant détaillé dans la section suivante.

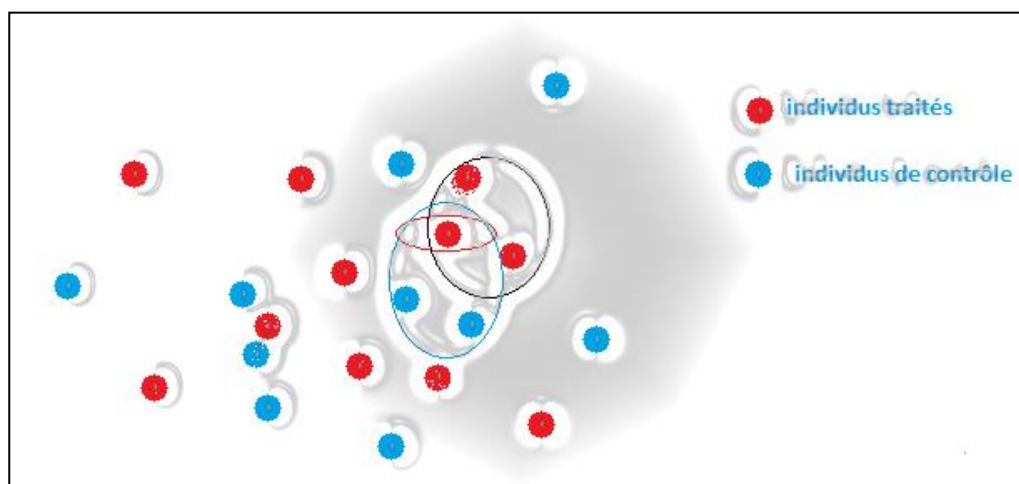
1.4. Intégration des effets d'interaction dans l'estimation de l'effet de traitement par l'économétrie spatiale

Pour prendre en compte les interactions, aussi bien dans le processus de sélection que dans l'estimation des équations de résultats et des effets de la politique, nous proposons un cadre d'analyse général des politiques non expérimentales. Cette proposition constitue une extension des méthodes d'estimation traditionnelles de l'effet de traitement. Elle permet, d'une part, de tenir compte des différentes formes d'interactions et d'autre part, de montrer, sous certaines conditions, que les estimations sous les hypothèses SUTVA et PIA en sont des cas particuliers.

Pour cela, nous supposons que l'ensemble des interactions sont générées à travers un processus spatial. Nous combinons les méthodes d'estimation standard de l'effet causal avec celles de l'économétrie spatiale. Ces dernières permettent d'intégrer dans les estimations les interactions liées aux caractéristiques observables, aux caractéristiques inobservables et aux interactions liées aux variables d'intérêt, aussi bien dans le processus de sélection que dans la performance de la politique.

Les types d'interactions que nous considérons sont les suivantes. On considère une population composée de deux types d'individus : des individus traités, constituant le groupe traité et des individus non traités dont certains peuvent servir d'individus de contrôle. Considérons l'individu traité de référence entouré en rouge (cf. Figure 1-3). Son voisinage, mixte, est constitué par l'ensemble des individus qui sont entourés en bleu et noir. Ce voisinage peut être décomposé en un « *voisinage traité* », un « *voisinage non traité* ». En toute généralité, on suppose que l'individu peut interagir différemment selon le type de voisinage.

Figure 1-3 : Les différentes formes d'interactions



Source : Sileymane BA-2015

Par exemple, dans l'évaluation d'une politique d'allègement fiscal pour les entreprises, nous pouvons avoir trois types d'effets liés à l'hétérogénéité du voisinage : les effets de compétition entre les communes bénéficiaires, les effets de compétition entre les communes non bénéficiaires et les effets de compétition entre les communes

bénéficiaires avec les communes non bénéficiaires. Dans le premier cas, *toutes choses égales par ailleurs*, chaque commune traitée va agir sur d'autres caractéristiques pour capitaliser au maximum sur les avantages du dispositif (attirer des nouvelles entreprises, élargir l'assiette fiscale,..). Dans les deux derniers cas, les communes non bénéficiaires vont agir sur d'autres caractéristiques, soit pour ne pas subir les conséquences du dispositif (délocalisation des entreprises, perte d'emplois,...), soit pour tirer des avantages du dispositif grâce à leurs proximités avec les communes traitées (création d'emploi, accroissement de l'assiette fiscale) (Givord et al., 2013; Rathelot et Sillard, 2008).

Comme dans la plupart des politiques non expérimentales, l'estimation de l'effet du dispositif dans le cadre du modèle canonique de Rubin (MCR) sous l'hypothèse SUTVA et PIA ne permet pas de prendre en compte l'effet réel du dispositif. La première hypothèse ignore les externalités engendrées par le dispositif et la seconde ne considère qu'une partie des effets d'externalités puisqu'elle ignore les effets entre les individus de statuts différents. Ainsi, avec les deux méthodes d'estimation sous SUTVA et PIA, l'effet de traitement estimé est biaisé (surestimé ou sous-estimé) et par conséquent, la prudence est de mise par rapport à toute décision politique à prendre à partir de ces résultats.

Partant des interactions potentielles entre les individus dans le processus de sélection et dans les résultats potentiels, nous présentons le modèle spatial de sélection (section 1.4.1), puis une spécification spatiale des équations de résultats potentiels (section 1.4.2)

1.4.1. Intégration des effets d'interaction dans le processus de sélection

Pour intégrer les effets d'interaction liés aux contextes (institutionnel ou géographique), nous supposons que la décision d'un individu de participer à une politique dépend certes de ses propres caractéristiques mais qu'elle peut également être influencée par les différentes formes d'interactions *à la Manski* : interactions endogènes, interactions exogènes, effets corrélés. Par conséquent, nous faisons face à des situations d'interdépendance des décisions des individus.

Ces effets doivent être considérés comme des éléments explicatifs de la décision de participation d'un individu à la politique et ne doivent pas être ignorés dans l'estimation du processus de sélection.

La problématique des interdépendances des décisions, choix ou encore préférences des individus n'est pas nouvelle et est largement traitée dans la littérature. Certains auteurs comme Brock et Durlauf, (2001), Cont et Löwe (2010), An (2005), Lee et *al.* (2010), Bramoullé et *al.*(2009) ou encore Liu et Lee (2010), abordent ces interdépendances sous la forme d'interactions sociales. Dans d'autres études, (Smirnov et Egan, 2012; Smirnov, 2010a, 2010b; Case, 1992; Ghosh, 2011; Goetzke, 2008 ; Fleming, 2004), les interdépendances sont traitées sous la forme d'interactions spatiales. Ces dernières peuvent être, cependant, une manière de formaliser les interaction sociales, les effets de pairs notamment (Calvó-Armengol et *al.*, 2009, Campos et Pinto, 2010; Graham, 2011; An, 2005).

Nous développons dans les paragraphes suivants le cas des interdépendances des décisions dues aux effets endogènes et aux effets corrélés. Ces deux types d'effets échappent au contrôle du chercheur.

1.4.1.1 Les interdépendances des décisions de participation des individus

On suppose que la décision d'un individu est fonction de ses propres caractéristiques mais aussi des décisions prises dans son voisinage. Cette situation fait référence aux interactions endogènes de Manski (1993).

Soit D_i une variable binaire indiquant la décision observée de l'individu. $D_i = 1$ signifie que l'individu participe à la politique et $D_i = 0$ indique que l'individu ne participe pas à la politique. Sous forme matricielle, l'équation de sélection prenant en compte la décision dans le voisinage, est définie comme suit :

$$D = \rho W D + Z \theta + \varepsilon_D \quad (1-58)$$

Où D est le vecteur ($N \times 1$) des observations de la variable dépendante, Z est la matrice ($N \times p$) des p variables dépendantes.

W représente la matrice de voisinage de dimensions ($N \times N$). Les interactions existantes entre les individus i et j sont évaluées par les termes de la matrice. Par définition les éléments w_{ii} de la diagonale de la matrice sont nuls, car un individu n'interagit pas avec lui-même. Les éléments hors diagonale de la matrice w_{ij} indiquent l'intensité de la proximité de l'individu i par rapport à l'individu j et exprime par le fait

l'intensité de l'interaction entre eux. θ est le vecteur $(p \times 1)$ des paramètres de régression.

ε_D est le vecteur $(N \times 1)$ des termes d'erreurs de la régression, suivant une loi normale, d'espérance nulle et de matrice des variances-covariances : $\varepsilon_D \sim N(0, \sigma^2 I)$.

ρ est le paramètre d'autocorrélation spatiale ou d'interaction, dont l'estimation permet d'évaluer l'influence des interactions existantes dans les prises de décisions D.

Pour tester la présence ou non des interférences entre les décisions, nous posons une hypothèse nulle d'absence d'interdépendance dans les décisions des individus.

Proposition 1 : Si le paramètre ρ est non significatif ($H_0 : \rho = 0$), cela signifie une absence d'interdépendances dans les décisions de participation à la politique, chaque individu prenant de manière indépendante sa décision. L'hypothèse SUTVA est postulée à raison.

En revanche, si le paramètre ρ est significatif et différent de 0, il y a interdépendances des décisions prises par les individus. Pour $\rho > 0$ alors la décision prise par un individu est similaire aux décisions prises par son voisinage. En revanche, si $\rho < 0$, la décision prise par un individu est opposée aux décisions prises par son voisinage. L'hypothèse SUTVA est postulée à tort.

1.4.1.2 Les interdépendances des décisions par le biais des interactions entre les inobservables

On suppose maintenant que la décision d'un individu est fonction de ses propres caractéristiques mais aussi des caractéristiques inobservables de son voisinage. Cette supposition amène à considérer que des facteurs inobservables peuvent affecter à la fois le traitement et les performances de la politique. Elle invalide de facto l'hypothèse CIA.

Elle fait référence aux *effets corrélés* de Manski (1993). Sous forme matricielle, l'équation de sélection est définie de la manière suivante :

$$D = Z\theta + \varepsilon_D ; \text{ avec } \varepsilon_D = \lambda W\varepsilon_D + e \quad (1-59)$$

$D, Z, W, \theta, \varepsilon_D$ sont définis comme précédemment. Le paramètre λ indique l'intensité de l'autocorrélation spatiale entre les résidus de la régression. e est le terme d'erreur suivant une loi normale, d'espérance nulle et de matrice des variances-covariances : $e \sim N(0, \sigma^2 I)$.

Si la matrice $(I - \lambda W)$ est non-singulière, le modèle se réécrit sous la forme suivante :

$$D = Z \theta + (I - \lambda W)^{-1} e \quad (1-60)$$

Proposition 2 : Si le paramètre λ est non significatif ($H_0 : \lambda = 0$), alors il n'y a pas d'interférences entre les inobservables de l'individu avec celles de son voisinage. En d'autres termes, il n'y a pas de facteurs inobservables qui influencent les décisions et par conséquent, la prise de décision de participation est indépendante de celles des autres individus. L'hypothèse CIA est postulée à raison.

En revanche, si le paramètre λ est significatif et différent de 0, alors il y a des interférences dans les inobservables de l'individu et celles de son voisinage. Cela signifie qu'un choc aléatoire observé chez un individu affecte non seulement la décision de cet individu mais aussi la décision des autres individus de son voisinage à travers l'opérateur $(I - \lambda W)^{-1}$. L'hypothèse CIA est postulée à tort.

En considérant ce qui précède, la significativité des paramètres spatiaux ρ (équation (1-58)) et λ (équation (1-60)) révèle les interdépendances des décisions des individus. Cela conduit par conséquent, à redéfinir l'équation de sélection dans les méthodes de *matching* (cf. 1.2.2) et dans l'estimation de l'effet de traitement en deux étapes à la Heckman (cf. 1.2.4)

En effet, dans le cadre des estimations par le *propensity score matching*, la significativité des paramètres spatiaux affecte (i) la qualité des estimateurs du modèle de sélection servant à construire le score de propension et (ii) met en cause la fiabilité du groupe de contrôle.

Dans le cadre d'une estimation à deux étapes à la Heckman, la significativité des paramètres affecte (i) la qualité des estimateurs du modèle de sélection servant à calculer

l'Inverse du Ratio de Mills et (ii) met en cause la fiabilité du contrôle du biais de sélection sur les inobservables.

Lorsque le paramètre ρ est significativement différent de 0, l'estimation du modèle de sélection par la méthode des moindres carrés ordinaires (MCO) conduit à des estimateurs biaisés et non convergents. En conséquence, les scores de propension servant à construire le groupe de contrôle avec les méthodes quasi-expérimentales et l'Inverse du Ratio de Mills servant à contrôler le biais de sélection dans une estimation en deux étapes, sont biaisés. *Toutes choses égales par ailleurs*, les estimateurs des paramètres d'intérêt (ATE, ATT, ATU) obtenus à partir de l'estimation des équations de sélection sont par conséquent biaisés.

Lorsque le paramètre λ est significativement différent de 0, l'estimation du modèle de sélection par la méthode des MCO conduit à des estimateurs non biaisés, mais inefficients. Par conséquent, les estimateurs servant à calculer l'Inverse de Ratio de Mills sont non biaisés, mais inefficients. *Toutes choses égales par ailleurs*, les estimateurs des paramètres d'intérêt (ATE, ATT, ATU) obtenus à partir de ces estimateurs sont par conséquent biaisés. Il faut noter lorsque le paramètre λ est significatif, cela implique la violation de l'hypothèse CIA et par conséquent les méthodes quasi expérimentales reposant sur cette hypothèse ne sont plus valides.

C'est pourquoi, pour une meilleure estimation des paramètres d'intérêt avec prise en compte des effets d'interaction ou interdépendances, il faut définir un nouveau score de propension spatial (PSMS) dans le cadre d'un *matching* et l'équivalent de l'Inverse de Ratio de Mills spatialisé (IRMS) dans l'estimation à *la Heckman*.

Il faut signaler que la prise en compte des interactions introduit de l'hétéroscédasticité dans les erreurs dans les équations de sélection (1-58) et (1-60), par conséquent le maximum de vraisemblance traditionnel n'est plus adapté pour l'estimation des paramètres. Ces équations de sélection peuvent être estimées par d'autres méthodes alternatives parmi lesquelles : la méthode *expectation maximization* (EM) de McMillen (1992, 1995), les méthodes *Generalized method of moments* (GMM) de Pinkse et Slade (1998) et enfin les méthodes bayésiennes développées par Lesage (1999, 2004).

1.4.2. Les effets d'interaction dans la performance de la politique

En considérant les effets d'interaction, les résultats potentiels des individus deviennent plus complexes. On suppose que le résultat observé pour un individu i est une fonction linéaire de ses caractéristiques observables, des effets des interactions et des inobservables.

Pour intégrer ces interactions dans les spécifications économétriques, nous supposons qu'elles suivent un processus de dépendance spatiale. Le résultat observé peut s'écrire sous la forme d'un modèle *General Nesting Spatial* (Tableau 1-1) :

$$Y = \rho WY + X\beta + WX\theta + \varepsilon ; \varepsilon = \lambda W\varepsilon + u \quad (1-61)$$

Y représente le vecteur du résultat observé (taux de réussite scolaire, taux d'insertion professionnelle, ...), X représente la matrice des variables explicatives, u le vecteur des termes d'erreurs avec les propriétés usuelles.

β comme θ représentent les vecteurs des paramètres associés respectivement aux matrices X et WX .

W représente la matrice de voisinage. Les paramètres ρ et λ mesurent l'intensité de la dépendance entre les individus et capturent respectivement les effets endogènes et les effets corrélés.

A partir de cette spécification (1-61), nous pouvons estimer la performance de la politique en fonction des différents types d'effets définis. Cependant, nous allons considérer dans la suite uniquement les effets endogènes dans l'estimation de la performance de la politique pour deux raisons principales. D'une part, la spécification *spatial autoregressive model* (SAR) permet de tester l'interdépendance des résultats entre un individu et son voisinage donc de vérifier indirectement l'hypothèse selon laquelle, le traitement reçu par un individu affecte son résultat mais aussi les résultats des autres individus. D'autre part, l'effet de traitement obtenu à partir de cette spécification est comparable avec ceux obtenus sous SUTVA et PIA, car dans les trois cas les hypothèses portent sur les interférences des résultats.

Par ailleurs, la spécification SAR permet de prendre en compte les effets des interactions mutuelles à travers le «*spatial multiplier*», comme le souligne, Soetevent (2006, p.199) : «*the endogenous and exogenous effects represent real social effects only the endogenous effect creates a feedback loop*».

En revanche, lors de la mise en œuvre de la politique, le voisinage devient hétérogène en ce sens que certains individus sont traités et d'autres non. Ainsi, les effets des interactions entre un individu selon son statut et son voisinage peuvent différer selon le type de voisinage considéré. C'est pourquoi, on suppose que les interactions entre un individu et son voisinage sont bidirectionnelles, c'est-à-dire qu'ils s'influencent mutuellement et de la même manière.

Dans ce cadre, le résultat potentiel d'un individu devient plus complexe, il résulte des interactions avec son voisinage traité et celles de son voisinage non traité, en plus de ses propres caractéristiques. Le résultat observé est donné sous forme matricielle par :

$$Y = \rho_{k1} W_k^1 Y + \rho_{k0} W_k^0 Y + X\beta_k + \varepsilon_k \quad (1-62)$$

Où Y représente la variable dépendante étudiée qui indique la performance de la politique (taux de réussite scolaire, taux d'insertion professionnelle, nombre le logement sociaux construits...).

Si $k = 1$ on obtient l'équation de résultat pour les individus traités et si $k = 0$ on a l'équation de résultat pour les individus non traités (groupe de contrôle). X représente la matrice de variables explicatives, ε_k le vecteur des termes d'erreurs suivant les hypothèses usuelles et β_k le vecteur des paramètres associés à X. Les matrices W_k^1 et W_k^0 de dimension $(N \times N)$ permettent de spécifier les quatre formes de voisinages pouvant exister à savoir les interactions intra groupes et inter groupes.

Si la matrice $(I_n - \rho_{k1} W_k^1 - \rho_{k0} W_k^0)^{-1}$ est inversible alors on peut écrire les résultats (équation(1-62) sous la forme réduite suivante :

$$Y = (I_n - \rho_{k1} W_k^1 - \rho_{k0} W_k^0)^{-1} X\beta_k + (I_n - \rho_{k1} W_k^1 - \rho_{k0} W_k^0)^{-1} \varepsilon_k \quad (1-63)$$

Où ρ_{k1} et ρ_{k0} , représentent les paramètres d'autocorrélation indiquant respectivement l'ampleur des interactions intra et intergroupes³.

La particularité de ces spécifications est que les effets des interactions endogènes sont modélisés pour explicitement prendre en compte les effets de diffusion de la politique.

A partir de la spécification de ce cas général de l'équation de résultat (équation (1-62)), on peut en tirer les cas particuliers sous lesquels les hypothèses SUTVA et CIA sont vérifiées.

Pour des raisons pratiques et de simplification des écritures, on considère uniquement le cas des individus traités mais cela est valable également pour les individus non traités (Figure 1-4). Dans ce cas, on a $k = 1$, donc ρ_{11} et ρ_{10} représentant respectivement le paramètre d'interaction intragroupe (un individu traité avec son voisinage traité) et le paramètre d'interaction intergroupe (un individu traité avec un voisinage non-traité).

Proposition 3 : Si $\rho_{11} = 0$ et $\rho_{10} = 0$, alors il y a absence d'interférences entre le résultat d'un individu avec son voisinage. En d'autres termes, le traitement reçu par un individu n'affecte pas les résultats des autres individus de son voisinage. On se retrouve alors le cas où l'hypothèse SUTVA est valide.

Proposition 4 : Si $\rho_{11} \neq 0$ et $\rho_{10} = 0$, alors il n'y a que des interférences intra groupe, c'est-à-dire qu'il y a des interactions entre l'individu traité avec son voisinage traité. Dans ce cas, l'hypothèse PIA (interférences intragroupe) est valide, c'est-à-dire qu'il y a des interactions à l'intérieur du groupe des traités.

Proposition 5 : Si $\rho_{11} = 0$ et $\rho_{10} \neq 0$, alors il n'y a que des interférences inter groupe, c'est-à-dire qu'il y a des interactions entre l'individu traité avec son voisinage non-traité. Cela implique une invalidation de l'hypothèse PIA.

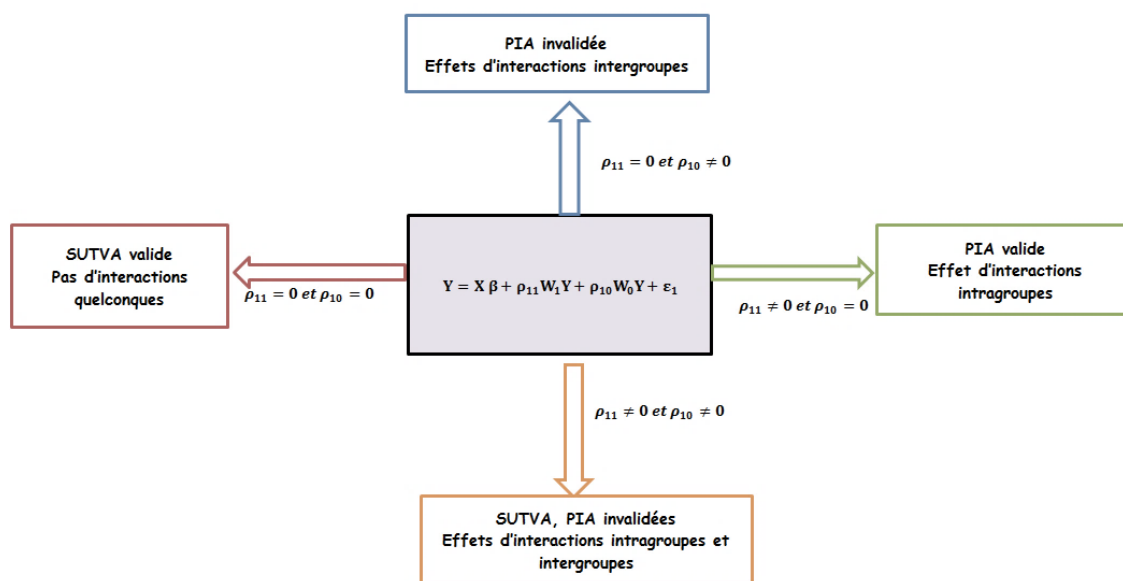
³ Cette partie est largement développée dans le Chapitre 4

Proposition 6 : Si $\rho_{11} \neq 0$ et $\rho_{10} \neq 0$, alors nous sommes en présence, à la fois d'interférences intra et intergroupes. Aucune des deux hypothèses SUTVA et CIA n'est valide.

Par ailleurs, lorsque les paramètres d'autocorrélation ρ_{k1} et ρ_{k0} dans l'équation (1-62) sont significativement différents de zéro, alors les estimations des paramètres des résultats (équation 1-62), par les MCO ne sont plus adaptées. L'estimation par la méthode des moindres carrés produit des paramètres β_k biaisés et non convergents.

Pour estimer l'équation de résultats (équation (1-62) avec deux matrices différentes, nous utilisons la méthode du maximum de vraisemblance développée par (Lacombe, 2004).

Figure 1-4 : L'approche spatiale dans l'évaluation des politiques publiques



Source : Sileymane Ba-2015

A la lumière de ce qui précède, nous allons redéfinir de nouveaux paramètres d'intérêt de l'effet de traitement tenant compte des effets d'interaction Plus précisément, nous introduisons le *spatial propensity score matching* lorsqu'on suppose que le biais de sélection est dû aux observables (cf. 1.4.2.1) et *l'Inverse du Ratio de Mills Spatialement ajusté* lorsqu'on considère que le biais de sélection est dû aux inobservables (cf. 1.4.2.2).

1.4.2.1 Identification des paramètres d'intérêt avec les effets d'interaction : le cas du *matching*

Spatial Propensity Score Matching SPSM

Il consiste à intégrer les effets d'interaction dans le calcul du *propensity score matching* (PSM) qui servira à construire le groupe de contrôle. Dans le cas du *spatial propensity score matching* (Chagas et al., 2012), on suppose que la probabilité d'un individu d'être traité, ou pas, P_{Spat} est fonction de ses propres caractéristiques mais aussi de la décision prise dans son voisinage (1–58).

$$P_{Spat}(x) = Pr(D = 1|X = x,) \in [0,1] = \Phi\{h(x), g(W * P_{Spat})\} \quad (1-64)$$

Φ représente la fonction de répartition de la loi normale, $h(x)$ est fonction des variables exogènes, $g(W * P_{SAR})$ la fonction des probabilités dans le voisinage et W représente la matrice de voisinage ou matrice de pondération spatiale.

Pour réaliser l'appariement en tenant compte des effets d'interaction, nous conservons le score de propension comme mesure de la distance entre deux individus. En conséquence, on obtient de nouveaux paramètres d'intérêt spatialement ajustés, qui tiennent compte des effets d'interaction entre les individus. Cette nouvelle distance est donnée par :

$$d_{ij}^{Spat} = |P_{Spat}(x_i) - P_{Spat}(x_j)| \quad (1-65)$$

Dans les spécifications énumérées dans la partie 1.2.2.2, on remplace la distance d_{ij} par la nouvelle distance d_{ij}^{Spat} , ce qui permet de redéfinir de nouveaux critères de similarités et de nouvelles pondérations $\omega_{ij}^{Spat} = f(d_{ij}^{Spat})$ pour faire le *matching*.

Estimation des paramètres d'intérêt avec prise en compte des interactions

En fonction du critère de similarité spatialisé retenu (plus proche voisin, k-plus proches voisins, *caliper* ou *radius matching*, *Kernel matching*), les estimateurs des paramètres spatialisés peuvent être réécrits en utilisant ω_{ij}^{Spat} , la pondération spatialisée

entre les individus i et j . On obtient alors les différents estimateurs spatialisés ΔATT_{Spat} , ΔATU_{Spat} et ΔATE_{Spat} .

- **L'estimateur de l'effet de traitement moyen sur les traités spatialisé** est donné par :

$$\Delta ATT_{Spat} = \frac{1}{N_1} \sum_{i \in \{D_i=1\}} \left(y_{1i} - \frac{1}{\sum_{j \in \{D_j=0\}} \omega_{ij}^{spat}} \sum_{j \in \{D_j=0\}} \omega_{ij}^{spat} y_{0j} \right) \quad (1-66)$$

- **L'estimateur de l'effet de traitement moyen sur les non traités spatialisé** est donné par :

$$\Delta ATT_{Spat} = \frac{1}{N_0} \sum_{i \in \{D_i=0\}} \left(\frac{1}{\sum_{j \in \{D_j=1\}} \omega_{ij}^{spat}} \sum_{j \in \{D_j=1\}} \omega_{ij}^{spat} y_{1j} - y_{0i} \right) \quad (1-67)$$

- **L'estimateur de l'effet moyen de traitement total** est donné par :

$$\Delta ATE_{Spat} = \frac{N_1}{N} \Delta ATT_{Spat} + \frac{N_0}{N} \Delta ATU_{Spat} \quad (1-68)$$

1.4.2.2 Les méthodes de régression avec prise en compte des interactions

Lorsqu'on considère les interactions dans l'estimation de l'effet causal avec les méthodes de régression, "*Before-After*", "*Différence-in-Différence*", les estimateurs des moindres carrés ordinaires sont biaisés et non convergents. C'est pourquoi en considérant les interactions, il faut utiliser la méthode du maximum de vraisemblance pour une estimation sans biais de l'effet causal.

Nous proposons dans le chapitre 3, une approche DID spatialisée, une extension détaillée de la méthode DID avec les effets d'interaction.

1.4.2.3 Evaluation en *DEUX ETAPES* avec prise en compte des interactions

Dans cette section, nous proposons une estimation en deux étapes à la Heckman (1979) en tenant compte des effets d'interaction (Flores-lagunes et Schnier, 2012 ; Ward et al. 2014).

En considérant les effets d'interaction, nous avons deux types d'équations de sélection. La première est un modèle de type SAR ((1-58) qui tient compte des effets endogènes que nous développons dans la section suivante. Le second est un modèle de type SEM (1-60) qui tient compte des effets corrélés.

Nous proposons dans le chapitre 4, une approche du modèle en deux étapes intégrant les effets corrélés dans l'équation de sélection et les effets endogènes et hétérogènes dans l'équation de résultats.

- *Spécification de l'équation de sélection de type SAR*

On suppose que la décision d'un individu dépend de ses propres caractéristiques mais aussi des décisions prises dans son voisinage :

$$\begin{aligned} D &= \rho_D W_D D + Z \theta + \varepsilon_D \\ &= (I - \rho_D W_D)^{-1} Z \theta + (I - \rho_D W_D)^{-1} \varepsilon_D \end{aligned} \quad (1-69)$$

Avec une variance

$$\text{Var}(v_D) = \sigma^2 \{(I - \rho_D W_D)'(I - \rho_D W_D)\}^{-1} \equiv \sigma_{v_D}^2$$

- *Spécification des équations de résultats avec les effets d'interaction*

En prenant en compte les effets d'interaction, les équations de résultats pour les individus traités et les non traités s'écrivent de la manière suivante.

Pour les individus traités, le résultat potentiel est égal à :

$$Y_1 = (I - \rho_1 W)^{-1} X \beta_1 + (I - \rho_1 W)^{-1} \varepsilon_1 \quad (1-70)$$

Avec une moyenne

$$E(Y_1|X) = (I - \rho_1 W)^{-1} X \beta_1$$

et une variance

$$\text{Var}(Y_1) = \sigma^2 \{(I - \rho_1 W)' (I - \rho_1 W)\}^{-1} \equiv \sigma_{v_1}^2$$

Pour les individus non traités, le résultat potentiel est égal à :

$$Y_0 = (I - \rho_0 W)^{-1} X \beta_0 + (I - \rho_0 W)^{-1} \varepsilon_0 \quad (1-71)$$

Avec une moyenne

$$E(Y_1|X) = (I - \rho_0 W)^{-1} X \beta_0$$

et une variance

$$\text{Var}(Y_0) = \sigma^2 \{(I - \rho_0 W)' (I - \rho_0 W)\}^{-1} \equiv \sigma_{v_0}^2$$

On suppose que la distribution jointe des termes (v_1, v_0, v_{eD}) des équations (1-69), (1-70) et (1-71) suit une loi normale de moyenne 0 et de matrices de variance-covariance :

$$\begin{pmatrix} v_D \\ \varepsilon_1 \\ \varepsilon_0 \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & r_{1v_D} \sigma_{v_1} & r_{0v_D} \sigma_{v_0} \\ r_{11} \sigma_{v_1} & \sigma_{v_1}^2 & r_{10} \sigma_{v_1} \\ r_{00} \sigma_{v_0} & r_{01} \sigma_{v_1} & \sigma_{v_0}^2 \end{pmatrix} \right]$$

- *Identification des paramètres d'intérêt*

A partir des équations (1-69), (1-70) et (1-71), on en déduit les trois paramètres d'intérêts.

On obtient pour **l'effet moyen du traitement sur les traités** intégrant les effets des interactions dans les équations de sélection et de résultats :

$$ATT_{SAR} = ((I - \rho_1 W)^{-1} X \beta_1 - (I - \rho_0 W)^{-1} X \beta_0) + [r_{1v_D} \sigma_{v_1} - r_{0v_D} \sigma_{v_0}] E \left[\frac{\phi(Z \theta^*)}{\Phi(Z \theta^*)} \right] \quad (1-72)$$

Le terme $\frac{\phi(Z\theta^*)}{\Phi(Z\theta^*)}$ représente l'Inverse du Ratio de Mills Spatialisé noté π_{1SAR} , estimé à partir de l'équation de sélection (1-69).

L'effet moyen du traitement sur les non traités intégrant les effets des interactions dans les équations de sélection et de résultats, est donné par :

$$ATU_{SAR} = \left((I - \rho_1 W)^{-1} X \beta_1 - (I - \rho_0 W)^{-1} X \beta_0 \right) + [r_{1v_D} \sigma_{v_1} - r_{0v_D} \sigma_{v_0}] E \left[\frac{\phi(Z\theta^*)}{1 - \Phi(Z\theta^*)} \right] \quad (1-73)$$

Le terme $\left[\frac{\phi(Z\theta^*)}{1 - \Phi(Z\theta^*)} \right]$ représente l'Inverse du Ratio de Mills Spatialisé noté π_{0SAR} , estimé à partir de l'équation de sélection (1-69).

Enfin **l'effet moyen du traitement sur la population** intégrant les effets des interactions dans les équations de sélection et de résultats, est égal à :

$$ATE_{SAR} = \left((I - \rho_1 W)^{-1} X \beta_1 - (I - \rho_0 W)^{-1} X \beta_0 \right) \quad (1-74)$$

A partir de l'équation (1-74), on récrit les équations (1-72) et (1-73) en fonction de l'effet ATE_{SAR} .

$$ATT_{SAR} = ATE_{SAR} + \pi_{1SAR} [r_{1v_D} \sigma_{v_1} - r_{0v_D} \sigma_{v_0}] \quad (1-75)$$

$$ATU_{SAR} = ATE_{SAR} + \pi_{0SAR} [r_{1v_D} \sigma_{v_1} - r_{0v_D} \sigma_{v_0}] \quad (1-76)$$

Les termes $\pi_{1SAR} [r_{1v_D} \sigma_{v_1} - r_{0v_D} \sigma_{v_0}]$ et $\pi_{0SAR} [r_{1v_D} \sigma_{v_1} - r_{0v_D} \sigma_{v_0}]$ représentent le biais de sélection spatialement ajustée dû aux inobservables.

Conclusion

L'objectif de ce premier chapitre était de proposer un cadre méthodologique pour la prise en compte des effets d'interaction sous forme de dépendance spatiale dans l'estimation de l'effet d'une politique.

Tout le long du chapitre nous avons montré le rôle des effets d'interaction, leurs sources et les mécanismes par lesquels ces effets se diffusent. Nous sommes revenus également sur les implications économétriques et économiques de l'évaluation des effets de traitement lorsque les interactions ne sont pas traitées.

Nous avons proposé une application des méthodes de l'économétrie spatiale afin de prendre en compte les effets d'interaction dans le processus de sélection et dans les équations de résultats. L'ensemble des outils proposés peut être testé dans la plupart des politiques publiques non expérimentales, comme les politiques territorialisées où l'existence des interactions entre les individus est potentiellement présente.

Les perspectives qui découlent de ce chapitre sont de proposer dans les 4 chapitres suivants un cadre théorique qui explicite les modélisations économétriques spatiales employées et leurs apports par rapport aux formalisations existantes. Nous proposons également d'illustrer ces nouvelles propositions par des applications empiriques.

Chapitre 2 Les interdépendances stratégiques dans les décisions de participation à une politique territoriale

Introduction

Dans son article premier, la loi n° 2010-597 du 3 juin 2010, relative au Grand Paris, définit le Grand Paris comme étant « *un projet urbain, social et économique d'intérêt national qui unit les grands territoires stratégiques de la région d'Ile-de-France, au premier rang desquels Paris et le cœur de l'agglomération parisienne, et promeut le développement économique durable, solidaire et créateur d'emplois de la région capitale. Il vise à réduire les déséquilibres sociaux, territoriaux et fiscaux au bénéfice de l'ensemble du territoire national. Les collectivités territoriales et les citoyens sont associés à l'élaboration et à la réalisation de ce projet* ».

Le projet du Grand Paris se développe pour et avec la construction d'une nouvelle infrastructure de transport en commun : le Grand Paris Express. Il vise également à favoriser « *la recherche, l'innovation et la valorisation industrielle* » et intègre finalement « *un objectif de croissance économique afin de soutenir la concurrence des autres métropoles mondiales* ». Pour atteindre ces différents objectifs, le projet s'appuie sur un ensemble de projets qui seront mis en œuvre dans certains territoires⁴ sur la base de Contrats de Développement Territorial (noté CDT dans la suite de la thèse) définis et réalisés conjointement par l'Etat, les communes et leurs groupements. Les dispositions relatives à la constitution et aux missions des CDT sont fixées au titre IV Développement territorial et projets d'aménagement, article 21, de la loi sur le Grand Paris précitée.

La constitution d'un Contrat de Développement Territorial « *porte sur le développement d'un territoire inclus dans un ensemble de communes d'un seul tenant et sans enclave* ». Il vise à mettre en place un cadre d'exécution et de mise en œuvre des stratégies économiques, sociales, politiques et environnementales qui accompagnent la construction du Grand Paris Express (GPE). C'est une démarche qui permet de prendre en compte à la fois les enjeux définis par l'État et les préoccupations des collectivités en matière de logements, d'aménagement, d'emploi, d'environnement ... qui sont perçues différemment selon les élus locaux. Pour exemple, les CDT participent à l'objectif du Grand Paris de construire chaque année 70 000 logements géographiquement et socialement adaptés en Ile-de-France sur la base d'un audit en leur sein leur permettant

⁴ Ces territoires ont été appelés selon les différentes formulations utilisées au cours de l'évolution du projet : *territoires de projet* ou *territoires stratégiques*.

d'identifier le nombre de logements à construire et le pourcentage de logements sociaux à réaliser.

Chaque CDT réunit plusieurs communes qui ont choisi d'y participer. De ce fait, les CDT sont élaborés de manière conjointe par l'Etat et les élus locaux. L'implication de l'Etat peut être perçue, d'une part, comme un gage de financement et de réalisation de certains projets d'aménagement (infrastructures de transport, désenclavement, construction de logements, création d'emploi,...). D'autre part, elle peut être perçue comme une perte d'autonomie des collectivités locales et une volonté de remettre en cause les principes de la décentralisation. En effet, si sur le principe, la participation à un CDT revient aux communes, elle est en partie « imposée » par la loi. L'article 7 de la loi n° 2010-597 du 3 juin 2010 relative au Grand Paris, confère à la Société du Grand Paris le pouvoir de se substituer à toute commune desservie par une nouvelle gare et non signataire d'un CDT, pour conduire des opérations d'aménagements autour des nouvelles gares du réseau de transport public du Grand Paris dans un rayon de 400m. Enfin, la région, le département concerné, l'association des maires d'Ile-de-France et le syndicat mixte Paris-Métropole sont consultés préalablement à la signature des contrats.⁵

C'est pourquoi, la participation d'une commune à CDT est le résultat d'un compromis ou d'un arrangement dans lequel les différents acteurs notamment les maires ont fait des renoncements, des concessions afin de s'accorder sur le CDT, sur les objectifs en matière de logement, l'emplacement de futures gares, Le CDT ne doit pas être perçu uniquement sous l'angle de compromis politiques, car il repose également sur des fondements économiques par lesquels les maires et les représentants de l'Etat se comportent comme de véritables agents économiques. Au final, le CDT doit être perçu comme le résultat d'un jeu d'acteurs économiques avec des interactions stratégiques entre les élus et les acteurs du Grand Paris.

L'objectif de ce chapitre est de proposer un cadre méthodologique d'analyse de la participation des communes franciliennes au projet du Grand Paris en intégrant les interdépendances stratégiques entre les communes concernées. Ces interactions

⁵ Article 21 de la loi n° 2010-597 du 3 juin 2010, relative au Grand Paris.

stratégiques peuvent être appréhendées à travers des interactions sociales et des interdépendances spatiales.

Manski (1993) classe les interactions sociales en trois types d'effets. Les effets endogènes qui représentent l'influence du comportement du groupe sur le comportement d'un individu. Les effets exogènes qui représentent l'effet des caractéristiques du groupe sur le comportement d'un individu et enfin les effets corrélés qui sont présents lorsque les individus agissent de manière identique parce qu'ils sont soumis à un même environnement.

Si l'importance des interactions sociales est un phénomène relativement bien connu, la prise en compte des interdépendances entre les individus dans l'évaluation des effets des politiques publiques – ou effets de traitement - reste encore peu développée. Pourtant, la structure et la portée de ces interactions sont de nature à enrichir la compréhension des résultats d'une politique (Ba et Baumont, 2015). En effet, une politique comprend un ensemble de mesures destinées à des individus localisés sur un territoire, qui appartiennent à des réseaux, à des groupes sociaux... et entretiennent entre eux des relations complexes. De ces relations naissent des interactions sociales, sous formes d'effets de voisinage, d'effets de pairs,..., qui peuvent influencer le comportement et l'attitude de chaque individu vis à vis de la politique. Par exemple, les interactions sont liées à la possibilité ou à la propension des individus à interagir, elles ne sont pas nécessairement symétriques, elles peuvent être hiérarchiques avec des effets de leader...

La formalisation des types d'interactions à *la Manski* (1993) peut être faite sous la forme de spécifications économétriques spatiales (cf. Tableau 1-1, Chapitre1). Par ailleurs la présence de ces effets d'interaction est susceptible d'affecter la qualité des estimations s'ils ne sont considérés. Ainsi, ne pas tenir compte de ces interactions positives (resp. négatives) revient à sous-estimer (resp. surestimer) les effets d'un programme (Affuso, 2010). En outre, prendre en compte ces interactions revient également à considérer des effets de diffusion, d'entraînement, des effets externes de la politique entre les individus.

Smirnov (2010a) indique que la dépendance spatiale implique l'interdépendance entre les individus et l'interdépendance entre leurs statuts. Par conséquent, « *ces interdépendances affectent les préférences des individus et créent un phénomène social qui influence la prise de décision, de sorte que les individus n'agissent pas de manière*

indépendante mais, aussi, ne prennent pas les décisions conjointement » (Smirnov, 2010a, p.3). Dans ce contexte, il s'agit bien d'analyser des interactions stratégiques et leurs effets dans les décisions des individus.

Un pan important de la littérature s'est développé en considérant les effets d'interaction dans les modèles de choix discrets binaires. Les travaux de Brock et Durlauf (2001; 2003; 2007), très largement repris dans la littérature, traitent les interdépendances sous forme d'interactions sociales et proposent une approche statistique pour traiter la question des interactions sociales dans les décisions des individus. D'une manière générale, dans la littérature, les interdépendances sont traitées sous trois formes : interactions sociales, effets de pairs, interdépendances spatiales. Plusieurs travaux s'intéressent plus précisément aux interdépendances spatiales.

Case (1992) considère les effets du voisinage, sous forme de dépendance spatiale, sur les décisions d'adoption d'une nouvelle technologie agricole par les paysans en Indonésie. Les résultats des estimations montrent que les effets de voisinage sont significatifs et positifs sur leurs décisions. Ces résultats soulignent que la non prise en compte des effets de voisinage peut conduire à une surestimation des effets des caractéristiques des ménages agricoles sur la probabilité d'adoption. S'intéressant aux choix de destinations touristiques (demande de loisirs), Smirnov et Egan (2012) analysent les interactions dans les choix des individus sous forme de dépendance spatiale inobservée (*effets corrélés*). Ils proposent un modèle d'utilité aléatoire spatialisé pour intégrer les interdépendances spatiales. Les estimations montrent que les interdépendances spatiales dans les choix des individus sont significatives et positives ce qui là encore revient à surestimer le rôle des caractéristiques des personnes dans les décisions qu'ils prennent.

Ghosh (2011) s'intéresse aux interactions stratégiques dans la décision des districts scolaires de participer au programme « *The Massachusetts Education Reform Act* ». La décision de participation des districts est modélisée en utilisant une forme spatiale du modèle probit. Les résultats montrent que la probabilité de participation d'un district est positivement affectée par les décisions de participation de ses voisins (*effets endogènes*). Ces résultats confirment de nouveau l'importance de ces formalisations pour détecter les interactions potentielles dans les processus de décision.

L'étude de Bandiera et Rasul (2006) analyse les décisions des agriculteurs d'adopter une nouvelle spéculation agricole par rapport aux choix de leurs réseaux familiaux ou amicaux. Les résultats montrent que les effets sociaux ne sont pas uniformes, car ils sont positifs lorsque le nombre d'adoptants est limité, et négatifs quand la nouvelle spéculation est adoptée par beaucoup de personnes du réseau. Les effets de réseaux puis les effets de compétition entre les agriculteurs qui s'accroît avec le nombre d'adoptants peuvent être avancés pour expliquer ces résultats.

Dans le cadre du Grand Paris, la prise en compte des interdépendances stratégiques dans les décisions des communes d'intégrer un CDT va être étudiée par les modèles de choix discret binaires et contribue à trois enjeux.

- Sur le plan économique, nous nous demandons si la décision de participation d'une commune est indépendante ou non des décisions des autres communes ? Si les décisions sont interdépendantes, qu'elle est alors la nature de ces interactions ? Les interactions positives traduisent un effet de complémentarité qui entraîne les communes proches à adopter le même comportement, tandis que les interactions négatives traduisent un effet de substituabilité qui va au contraire montrer que les communes proches adoptent des comportements différents. Les interactions entre les décideurs politiques locaux peuvent conduire à l'émergence d'un consensus autour des objectifs communs ou au contraire révéler des perceptions différentes.

- D'un point de vue économétrique, les estimations obtenues rendent compte de manière correcte les effets des caractéristiques des communes sur la décision.

- Pour les politiques territoriales, enfin, cette analyse permet d'éclairer la cohérence des choix des territoires stratégiques sélectionnés pour bénéficier des actions de développement dans le cadre du projet Grand Paris.

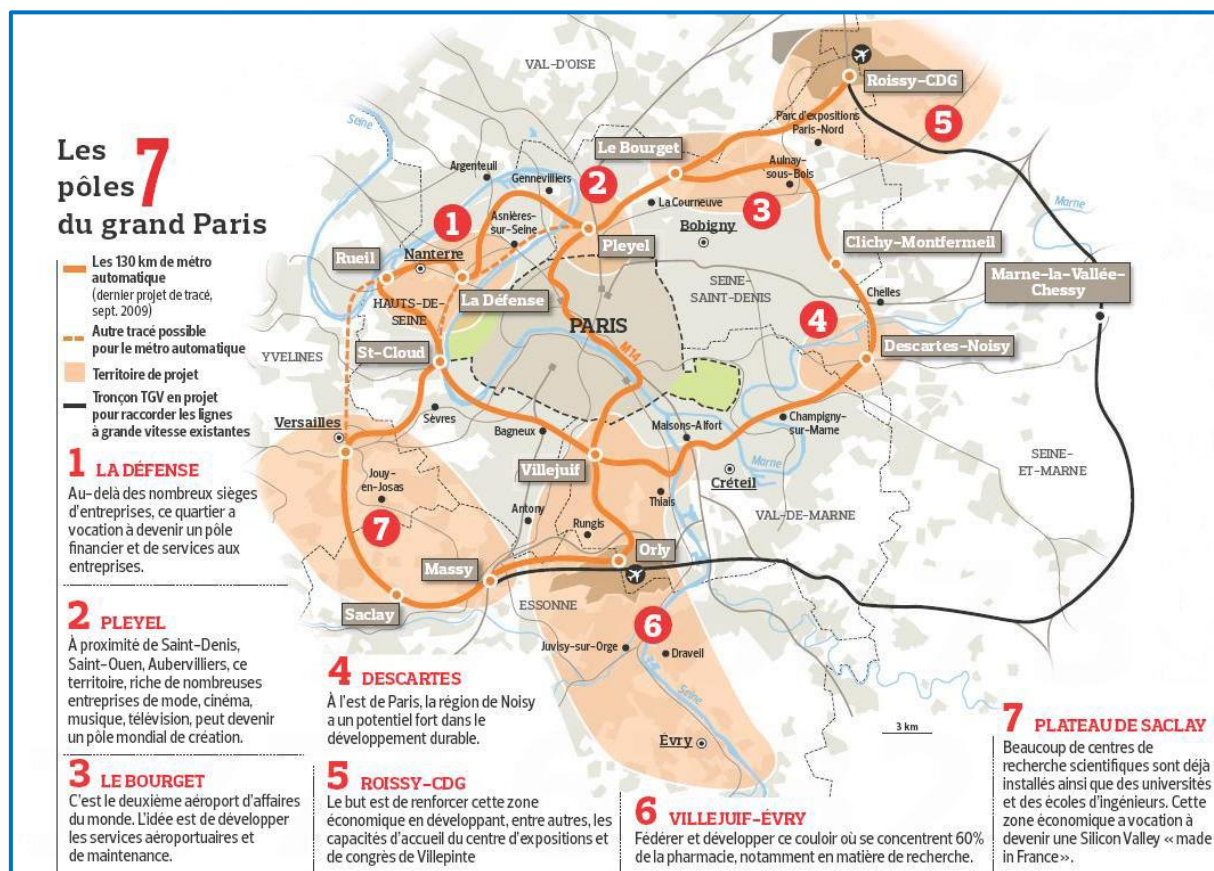
La suite de l'étude est structurée de la manière suivante. Nous présentons, dans la première partie (2.1), la politique du Grand Paris. Dans la deuxième partie (2.2), nous proposons une méthode d'estimation du processus de décision de participer ou non à une politique, en tenant compte des interdépendances stratégiques entre les individus. Cette proposition est appliquée, à travers le processus d'intégration des communes à un Contrat de Développement Territorial. Le périmètre d'étude et les données sont présentés dans la troisième partie (2.3) et les résultats sont donnés dans la quatrième et dernière partie (2.4).

2.1 Présentation et enjeux du projet Grand Paris

Le projet Grand Paris vise à dynamiser la région francilienne, en agissant sur un certain nombre de leviers économiques, afin d'améliorer l'attractivité et la compétitivité de la région (Prager, 2014). Conçu autour de la ville de Paris, il ambitionne de hisser la métropole Parisienne au rang des villes mondiales, à l'instar de Tokyo, Chicago, New York ou Londres. Pour cela, un ensemble de leviers ont été identifiés. L'action emblématique du projet concerne la construction du métro « Grand Paris Express » (Beaucire et Dauzelle, 2013)⁶ - noté GPE par la suite – avec un total de 70 gares dont 57 nouvelles, réparties sur un tracé de 200 km (*cf.* Carte 2-2). Ce projet vise, par ailleurs, un objectif de production de 70 000 logements par an (Orfeuil, 2014). Il s'appuie sur des effets de concentrations favorables à la croissance économique par l'identification de territoires stratégiques ou pôles économiques, répartis sur l'ensemble du territoire, mais dédiés à des activités spécialisés : la santé, la culture, la finance ... (*cf.* Carte2-1). Enfin, le projet du Grand Paris mise également sur le développement de la recherche et l'innovation, par le renforcement des pôles de compétitivité et par la valorisation du pôle scientifique et technologique du plateau de Saclay qui regroupera dans le futur un nombre très importants d'écoles et de centres de recherche. On estime que ce projet va générer, en moyenne, 53 000 à 67 000 emplois par an (Orfeuil, 2014). Le coût global du projet est estimé à 26 milliards d'euros à l'horizon 2030.

⁶ L'hypothèse, par ailleurs théoriquement fondée, est que cette infrastructure de transport va exercer une action structurante pour la région. Ses impacts font notamment l'objet de diverses études dans le cadre du conseil scientifique du Grand Paris. <http://www.societedugrandparis.fr/>

Carte 2-1 : Les 7 pôles du Grand Paris



Source : Le Figaro. <http://www.lefigaro.fr/assets/images/grand-paris.jpg>, page consultée le 14/09/2015

Cet ambitieux projet est institutionnalisé par la loi du 3 juin 2010, portant création de l'établissement public « Société du Grand Paris ». La Société du Grand Paris a pour mission « de concevoir et d'élaborer le schéma d'ensemble et les projets d'infrastructures composant le réseau de transport public du Grand Paris et d'en assurer la réalisation, qui comprend la construction des lignes, ouvrages et installations fixes, la construction et l'aménagement des gares, y compris d'interconnexion, ainsi que l'acquisition des matériels roulants conçus pour parcourir ces infrastructures et, dans les conditions de l'article 16, leur entretien et leur renouvellement, dans les conditions prévues par la présente loi »⁷.

⁷ JORF n°0128 du 5 juin 2010 page 10339, consultée le <http://www.legifrance.gouv.fr/>, 16/02/15.

L'articulation du projet Grand Paris autour du transport n'est pas fortuite, car la congestion reste un véritable problème dans la région d'Ile-de-France, que ce soit sur les réseaux routiers ou les transports en commun. Les résultats des enquêtes « ménages/déplacements » menées dans la région montrent que les problèmes concernent aussi bien les individus désirant rejoindre la capitale que les individus gravitant autour de cette dernière, naviguant de banlieues à banlieues. Le tracé du GPE correspond à deux boucles autour de la capitale qui relieraient les banlieues tout en passant par le centre de Paris (cf. Carte 2-2). A ces problèmes s'ajoute la question de l'isolement spatial, véritable frein au développement de certaines communes mal desservies.

Carte 2-2 : Schéma d'ensemble du Grand Paris



Source : <http://www.grand-paris.jll.fr/fr/le-projet-du-grand-paris/grand-paris-express/id/15>, consultée, le 17/09/2015

Ce tracé montre que le futur GPE desservira les zones pour le moment isolées, telles que le plateau de Saclay, mais aussi des zones enclavées par des éléments d'aménagement urbain tel qu'un aéroport ou encore une voie rapide comme une partie de

la Seine-Saint-Denis. Enfin, son tracé devrait permettre de relier plus rapidement les différentes gares TGV de la capitale avec les aéroports.

Par conséquent, l'amélioration de la mobilité conduit à faciliter les interactions entre les ménages et les entreprises, avec des effets attendus sur l'emploi et le chômage (L'Horty et Sari, 2013) et les interactions entre les activités économiques, avec des effets attendus sur l'innovation et la compétitivité des entreprises.

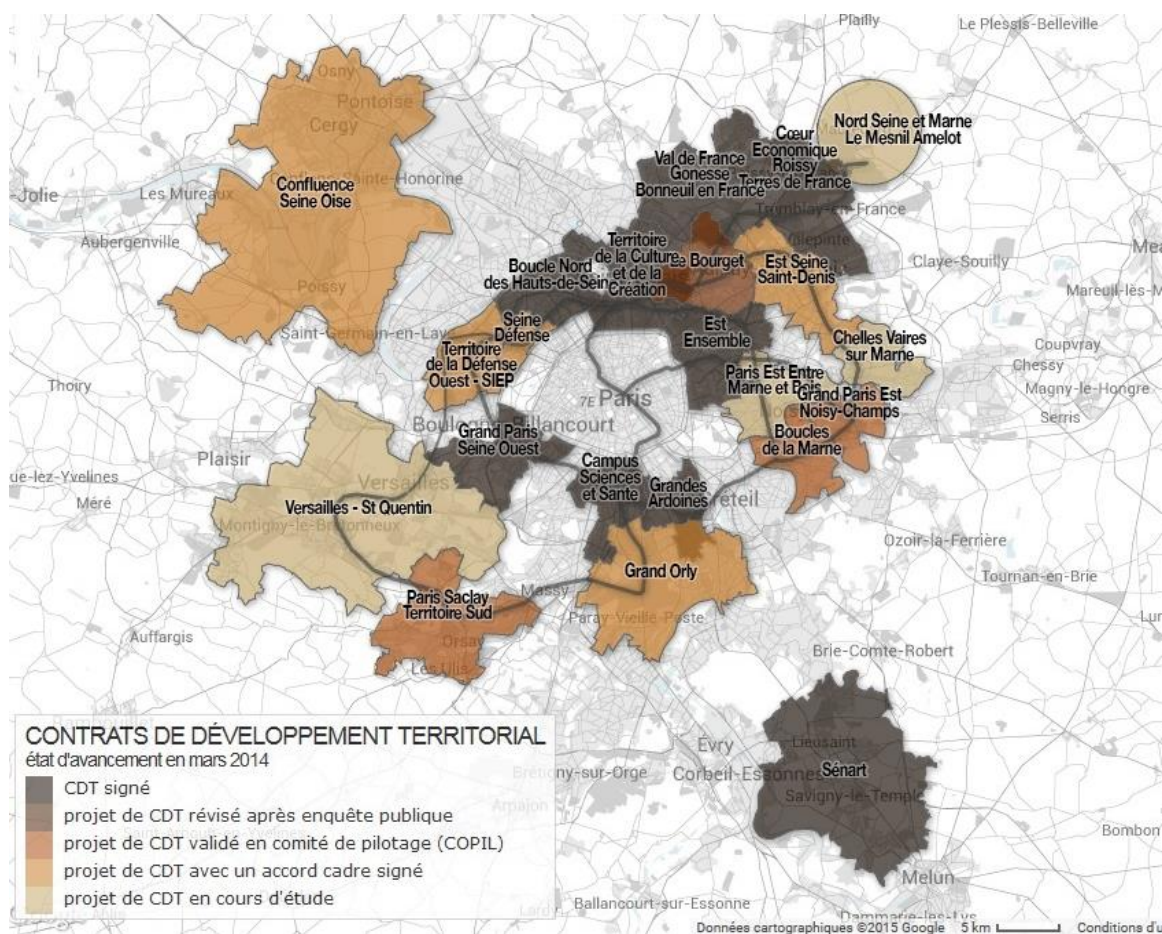
Le tracé du Grand Paris Express constitue dès lors un des paramètres de référence pour l'identification du territoire qui sera impacté par la politique d'aménagement (voir partie 2.3). L'organisation territoriale du Grand Paris est basée sur la mise en place de « territoires de projets » dont les axes et caractéristiques de développement font l'objet d'un Contrat de Développement Territorial (CDT) entre les communes concernées, l'Etat et la Région Ile-de-France. On dénombre actuellement 21 CDT (cf. 2.2) actifs, un 22^{ème} CDT restant encore à finaliser⁸. L'ensemble de ces CDT regroupe 156 communes ce qui représente 12 % des communes franciliennes et 8 % du territoire régional. L'ensemble de ces communes totalisent 4,4 millions d'habitants (38 % de la population francilienne) et 2,1 millions d'emplois (39 % des emplois franciliens)⁹.

La plupart des communes qui ont adhéré à un CDT se situent en petite couronne et en périphérie de la ville de Paris. Les différents CDT qui accompagnent les territoires du projet, insistent sur trois mécanismes de développement et d'attractivité : améliorer l'accessibilité via les infrastructures de transport, développer l'innovation en créant un climat de la connaissance et agir sur le cadre de vie (transports, habitat, environnement et culture). L'ampleur de ce dispositif interroge naturellement sur ses effets attendus.

⁸ L'ensemble des CDT signés peuvent être consultés sur le site de la Société du Grand Paris et à l'adresse suivante : <http://www.ile-de-france.gouv.fr/gdparis/CONTRATS-DE-DEVELOPPEMENT-TERRITORIAL2/CDT-signes-a-telecharger>.

⁹ <http://www.societedugrandparis.fr/projet/le-grand-paris/les-contrats-developpement-territorial>

Carte 2-3 : Situation des CDT – 1^{er} avril 2014



Source : <http://ateliergrandparis.fr/aigp/conseil/avis.php#carte>, consultée le 17/09/2015

La mise en œuvre d'une politique de cette ampleur interroge également les comportements des acteurs territoriaux franciliens compte tenu de la durée sur laquelle le projet s'étale. En effet, le périmètre du Grand Paris Express et le positionnement des territoires stratégiques couvrent environ les territoires de 500 communes, seules 156 d'entre elles se sont engagées dans des CDT.

Dans les chapitres précédents, nous avons souligné l'importance que peut prendre l'analyse du processus de sélection des individus impliqués dans la compréhension des effets d'une politique. Notre analyse consiste donc à étudier le processus d'adhésion des municipalités en introduisant comme facteur de décision les interactions entre les communes.

2.2 Modélisation de la décision de participation au Grand Paris

C'est en 2007 que germe le projet du Grand Paris pour une métropole du futur articulant projets urbanistiques et infrastructures de transports. Des projets, dessinés et imaginés sur les plans, aux projets retenus et en chantier, se passe toute une période, qui s'étale sur 7 années avec en 2014 l'identification quasi définitive des 22 CDT¹⁰ qui associeront les communes au projet. Au cours de cette période, la construction institutionnelle du projet est active et marquée par quelques grandes étapes et dates : les ateliers du Grand Paris (octobre 2009), la promulgation de la loi pour le Grand Paris (3 juin 2010) et la création de la Société du Grand Paris (juin 2010) et le décret relatif aux CDT (24 juin 2010), la signature de plusieurs CDT (entre octobre 2013 et février 2014), la promulgation de la loi de modernisation de l'action publique territoriale et d'affirmation des métropoles (MAPAM) en janvier 2014. Le planning prévisionnel de la construction du métro Grand Paris Express formalise par ailleurs les étapes de sa construction, rendant de plus en plus concret le projet et ses impacts ... jusqu'à la date finale prévue en 2030 (cf. Carte 2-2).

Au cours de la période, plusieurs débats et réunions publiques ont été menés, des campagnes d'informations « sur mesure » ont été effectuées, un site Internet a été mis à la disposition du public. Ces échanges ont porté sur différentes thématiques allant de l'information sur le projet en tant que tel, jusqu'aux tarifications, en passant par le calendrier des aménagements et de la construction du GPE, les coûts, les impacts environnementaux, socioéconomiques, fonciers... Ces actions de communication et de sensibilisation sont susceptibles d'affecter les comportements des décideurs locaux vis-à-vis du projet.

2.2.1 L'interdépendance des décisions des communes

On suppose que les maires des communes se comportent en agents économiques rationnels : chaque maire a comme objectif de maximiser son profit en captant, pour sa commune, le maximum de bénéfices engendrés par le Grand Paris. Cet objectif peut être, par exemple, l'accroissement de l'assiette fiscale de sa commune qui viendrait de

¹⁰ Un CDT sur les 22 n'a pas encore finalisé son périmètre en 2014.

l'accroissement de la valeur du parc résidentiel, du développement de ce parc résidentiel et de l'activité économique. Par conséquent, pour atteindre ses objectifs, chaque commune va mettre en place des dispositifs attractifs pour attirer des entreprises, améliorer le cadre de vie des habitants, développer le tissu économique et résidentiel. Elle peut engager des dépenses publiques pour l'amélioration des services (crèches, écoles, loisirs ...), la réalisation d'aménagements sectoriels (immobilier d'entreprises, parkings, voiries,...) ou bien encore la construction de logements sociaux...

Dans ce cadre, le projet Grand Paris peut engendrer une compétition fiscale ou un mimétisme fiscal entre les communes (Case et al., 1993 ; Feld et al., 2002; Leprince et al., 2005; Solé-Ollé, 2006) . Cela signifie que les décisions d'une commune ne sont pas indépendantes des décisions des autres communes voisines, comportements qui ont été formalisés et étudiés par tout un ensemble de travaux (Brueckner, 1998, 2003 ; Brueckner et Saavedra, 2001 ; Jensen et Toma, 1991). Chaque commune, dans sa prise de décision, intègre les décisions des autres communes voisines si bien que les externalités horizontales sont considérées dans ses choix. Les effets de débordements associés aux équipements sont notamment bien connus : par exemple la construction d'une infrastructure sportive (par exemple une piscine) ou culturelle (par exemple, un musée) ne profitera pas uniquement aux habitants de la commune mais elle profitera aussi aux populations des communes voisines.

Pour les décisions des communes franciliennes vis-à-vis des CDT, cela revient à se demander si les retombées attendues du projet ne pourraient pas être captées, quel que soit le degré d'implication de la commune ? Ici ce degré d'implication est binaire : soit la commune participe à la politique en intégrant un CDT, soit elle ne participe pas à la politique en décidant de rester en dehors des CDT.

La prise en compte des interactions dans les décisions des individus a été faite dans plusieurs travaux à travers des modèles de choix discrets et est formalisée par un modèle probit spatial (Ba et Baumont, 2015 ; Case, 1992b ; Ghosh, 2011 ; Smirnov, 2010a ; Smirnov et Egan, 2012).

On considère l'ensemble des communes de notre périmètre d'étude comme un ensemble fini d'individus noté N . Un CDT se constitue par l'adhésion de communes i et

on suppose que la constitution des CDT résulte des choix stratégiques des communes. Chaque commune décide d'intégrer, ou non, un CDT en fonction de ses propres caractéristiques, mais aussi en fonction de la décision des communes voisines. Suivant le modèle proposé par Brueckner (2003), chaque commune i est caractérisée par une fonction de résultats U définie comme suit :

$$U_i = (d_i, d_{-i}, X_i) \forall i, j \in N \quad (2-1)$$

d_i est la décision de la commune i de participer ou de ne pas participer (i.e. d'intégrer ou de ne pas intégrer un CDT). d_{-i} représente le vecteur des décisions prises par les autres communes j autres que la commune i . x_i est un vecteur de caractéristiques observables de la commune i .

A l'équilibre de Nash, chaque commune i choisit sa meilleure réponse, c'est-à-dire la stratégie d_i qui maximise sa fonction de résultats (2-1), les autres communes en faisant autant. La décision d_i^* solution pour la commune i , est obtenue par :

$$d_i^* = \tau_i(d_{-i}; x_i) \quad (2-2)$$

τ_i représente la fonction des meilleures réponses de i aux décisions des autres communes j .

L'équilibre de Nash est atteint lorsque chaque commune joue sa meilleure réponse :

$$d_i^* = \tau_i(d_{-i}^*; x_i) \forall i \in N, \forall d_i \neq d_i^* \quad (2-3)$$

L'étude empirique de l'existence d'interactions stratégiques se résume généralement à un test de significativité du coefficient τ_i de la fonction de réaction (Brueckner, 2003). On parle d'interactions stratégiques lorsque le coefficient de la fonction de réaction est significativement différent de 0. En d'autres termes, si l'hypothèse nulle $H_0 : \tau_i = 0$ est rejetée, alors on en déduit la présence d'interactions stratégiques entre les décisions des communes.

2.2.2 Spécification du processus de décision

Concrètement, la procédure de construction des CDT s'est étalée dans le temps et s'est faite de manière progressive. L'adoption définitive des CDT, dans leurs versions actuelles, a démarré le 28 Octobre 2013¹¹, avec la signature du CDT « Campus sciences et santé ». Aujourd'hui, 9 CDT ont été signés, 5 CDT ont été validés, 4 CDT ont un accord-cadre signé, 3 projets de CDT sont en cours d'étude et le CDT (Grand territoire Roissy) est en cours de définition¹². Cette temporalité nous amène à supposer qu'il existe un processus sous-jacent d'intégration des communes aux CDT, qui est marqué par l'observation des comportements des autres communes. Leurs propres décisions peuvent être expliquées par différents facteurs internes aux communes et relatifs aux enjeux économiques, démographiques et politiques qu'elles peuvent poursuivre. La mise en place d'un projet d'envergure d'aménagement du territoire francilien n'est, en outre, pas singulière face à l'architecture institutionnelle déjà existante des Etablissements Publics de Coopération Intercommunale au sein de la région¹³.

Compte tenu du processus de constitution des CDT, nous supposons que la décision des communes d'intégrer un CDT est déterminée de manière conditionnelle aux décisions des communes voisines et nous supposons que l'élaboration des CDT suit un processus markovien. En d'autres termes, la décision d'un individu est déterminée par les décisions des autres individus situés dans son voisinage immédiat, *selon un processus markovien dans un cadre spatial* (Cressie et Chan, 1989).

Cette approche, popularisée par (Besag, 1974) repose sur le principe suivant : "*the conditional distribution at site i depends only upon the values at those sites which are, in some sense, in the proximity of site i* ". Elle permet d'analyser un phénomène qui apparaît dans un endroit et qui se propage immédiatement dans le voisinage de ce site. Dans le cadre de la modélisation d'une variable dépendante binaire comme la décision d'intégration ou non d'une commune à un CDT, nous considérons le modèle auto-

¹¹ Les CDT dans leurs versions actuelles sont issus des schémas de développement territorial entamés depuis 2009.

¹² <http://www.grand-paris.jll.fr/le-projet-du-grand-paris/les-contrats-de-developpement-territorial-%28cdt%29/id/51>, consultée, le 18/02/2015.

¹³ Les communes franciliennes font parties des Communautés d'Agglomération, des Communautés de Communes, ou de Syndicats d'Agglomération Nouvelle.

logistique (Besag, 1974, 1975; Cressie, 1993) qui est très largement utilisé en écologie, pour modéliser l'apparition d'une espèce végétale dans un endroit donné et sa propagation ou sa dispersion dans son voisinage. C'est un type de modèle qui s'avère approprié et est utilisé pour modéliser une variable binaire spatialement corrélée.

Pour notre étude de cas, le choix d'une spécification auto-logistique est par ailleurs principalement motivé par trois raisons. D'abord, la décision de la commune d'intégrer un CDT n'est pas indépendante de la décision des autres communes voisines. Ensuite, les décisions des communes de participer ou non ne sont pas prises de manière simultanée et enfin le processus de constitution des CDT, nous amène à supposer que l'adhésion d'une commune à un CDT est conditionnelle à la participation des autres communes voisines. Sur ce dernier point en effet, le CDT ne peut exister que si d'autres communes adhèrent au dispositif.

La composante spatiale du modèle auto-logistique permet, dans un premier temps, de prendre en compte les interdépendances des décisions des communes et, dans un second temps, de qualifier la nature des décisions. Celles-ci peuvent être similaires ou opposées. On s'attend en particulier, à ce que les communes voisines adoptent des décisions similaires aux communes qui ont déjà intégré un CDT.

La décision d'intégration de la commune à un CDT à travers un modèle auto-logistique est formalisée de la manière suivante (Griffith, 2004).

Soit l'ensemble N des communes et y_i^* un index inobservable qui mesure la probabilité d'intégrer ou non un CDT :

$$y_i^* = \beta_0 + x_i' \beta + \rho \sum_{j \neq i} w_{ij} y_j^* + \varepsilon_i \quad (2-4)$$

x_i est un vecteur des caractéristiques de la commune qui interviennent dans la décision d'intégrer ou non le CDT : caractéristiques socio-démographiques, équipements des communes, situation géographique..., β est le vecteur des paramètres associés aux variables explicatives x . ε_i est le terme d'erreurs *i.i.d* et $\mathcal{N}(0, \sigma^2)$.

Les termes w_{ij} sont les éléments de la matrice W , de dimension $(N \times N)$, formalisant les interactions spatiales entre les communes i et j . La composante spatiale du

modèle ($\sum_{j \neq i} w_{ij} y_j$) indique que la décision d'une commune i , est influencée par les décisions prises par les autres communes j voisines, en fonction, d'une part, de l'intensité de la proximité entre les communes et, d'autre part, d'un périmètre de voisinage au-delà duquel les interactions sont supposées être nulles. Par convention, les éléments w_{ii} de la diagonale sont nuls.

ρ est le paramètre spatial, mesurant l'interdépendance conditionnelle entre les décisions des communes.

L'index y_i^* n'est pas observable directement, mais se déduit, d'une part, de l'observation de la décision de la commune et, d'autre part, des facteurs explicatifs sous-jacents à la prise de décision. Pour chaque commune on observe une variable y_i , binaire qui renseigne sur l'intégration ou non de la commune à un CDT :

$$y_i = \begin{cases} 1, & \text{si la commune intègre un CDT} \\ 0, & \text{sinon} \end{cases} \quad (2-5)$$

En considérant l'équation (2-5) et en supposant que l'intégration d'une commune à un CDT est fonction de ses propres caractéristiques et elle est conditionnelle aux décisions des communes voisines, la probabilité de participation p_i est définie de la manière suivante :

$$p_i = P(y_i = 1 | x, y_j) = \frac{\exp(\beta_0 + x'_i \beta + \rho \sum_{j \neq i} w_{ij} y_j)}{1 + \exp(\beta_0 + x'_i \beta + \rho \sum_{j \neq i} w_{ij} y_j)} \quad (2-6)$$

L'estimation du paramètre spatial ρ donne lieu à plusieurs interprétations sur la nature des interactions stratégiques entre les communes.

- Si ρ est non significatif ou nul, cela veut dire qu'il y a une absence d'interaction spatiale dans le modèle. Donc l'équation (2-6) correspond à un modèle logistique standard. Ce premier constat implique que la décision prise par la commune i est indépendante des décisions des autres communes voisines.

- Si le paramètre ρ est significatif différent de 0, alors la décision prise par une commune, est dépendante des décisions prises dans le voisinage.

Si ρ est positif, les décisions des communes peuvent être perçues comme complémentaires. Les décisions prises sont similaires. Ainsi, une commune voisine d'autres communes, déjà, membres d'un CDT a une probabilité plus importante d'intégrer un CDT par rapport à une commune ayant dans son voisinage des communes hors d'un CDT.

Si ρ est négatif, les communes et leur voisines prennent des décisions divergentes. Cette situation peut traduire différentes stratégies. Des communes relativement entourées de CDT peuvent jouer une stratégie de *passager clandestin* en décidant de ne pas intégrer un CDT, ceci pour éviter les contraintes institutionnelles qu'une telle appartenance implique, tout en espérant bénéficier des effets de débordement positifs de son voisinage. Pour d'autres communes, la stratégie concurrentielle consiste simplement à se démarquer de ces voisines.

2.2.3 Estimation du modèle auto-logistique spatial

Guyon (2007) souligne que l'introduction de l'interdépendance spatiale dans l'équation (2-6) implique que la méthode du maximum de vraisemblance classique n'est plus adaptée pour estimer des paramètres $\Omega(\alpha, \beta, \rho)$ du fait des difficultés de calcul de la constante de normalisation dans la log vraisemblance du modèle auto-logistique. La fonction de vraisemblance de l'équation (2-6) est égale à:

$$l(\Omega(\alpha, \beta, \rho)) = \prod_{i=1}^N \left(\frac{\exp(\beta_0 + x_i' \beta + \rho \sum_{j \neq i} w_{ij} y_j)}{1 + \exp(\beta_0 + x_i' \beta + \rho \sum_{j \neq i} w_{ij} y_j)} \right)^{y_i} \left(\frac{1}{1 + \exp(\beta_0 + x_i' \beta + \rho \sum_{j \neq i} w_{ij} y_j)} \right)^{1-y_i}$$

La Log-Vraisemblance correspond à :

$$L(\Omega) = \log(l(\Omega)) = \sum_{i=1}^N y_i \ln \left(\beta_0 + x_i' \beta + \rho \sum_{j \neq i} w_{ij} y_j \right) - [1 - y_i] \ln \left(1 + \left(\beta_0 + x_i' \beta + \rho \sum_{j \neq i} w_{ij} y_j \right) \right)$$

Avec $\Omega(\alpha, \beta, \rho)$, le vecteur des paramètres à estimer.

Pour contourner les difficultés de calcul avec le maximum de vraisemblance traditionnel, deux méthodes alternatives ont été développées pour une meilleure estimation des paramètres $\Omega(\alpha, \beta, \rho)$: le pseudo maximum de vraisemblance (PML)

(Besag, 1974 ; Cressie, 1993) et les méthodes de Monte-Carlo par les Chaînes Markoviennes (MCMC) (He et *al.* , 2003 ; Sherman et *al.*, 2006).

Il faut noter que le pseudo maximum de vraisemblance est très pratique et facile à mettre en œuvre avec des estimateurs convergents (Griffith, 2004). Les estimateurs PML sont proches des estimateurs du maximum de vraisemblance traditionnel si la dépendance spatiale est faible, mais en général ils sont moins efficaces.

2.3 Présentation des données et des structures d'interactions spatiales

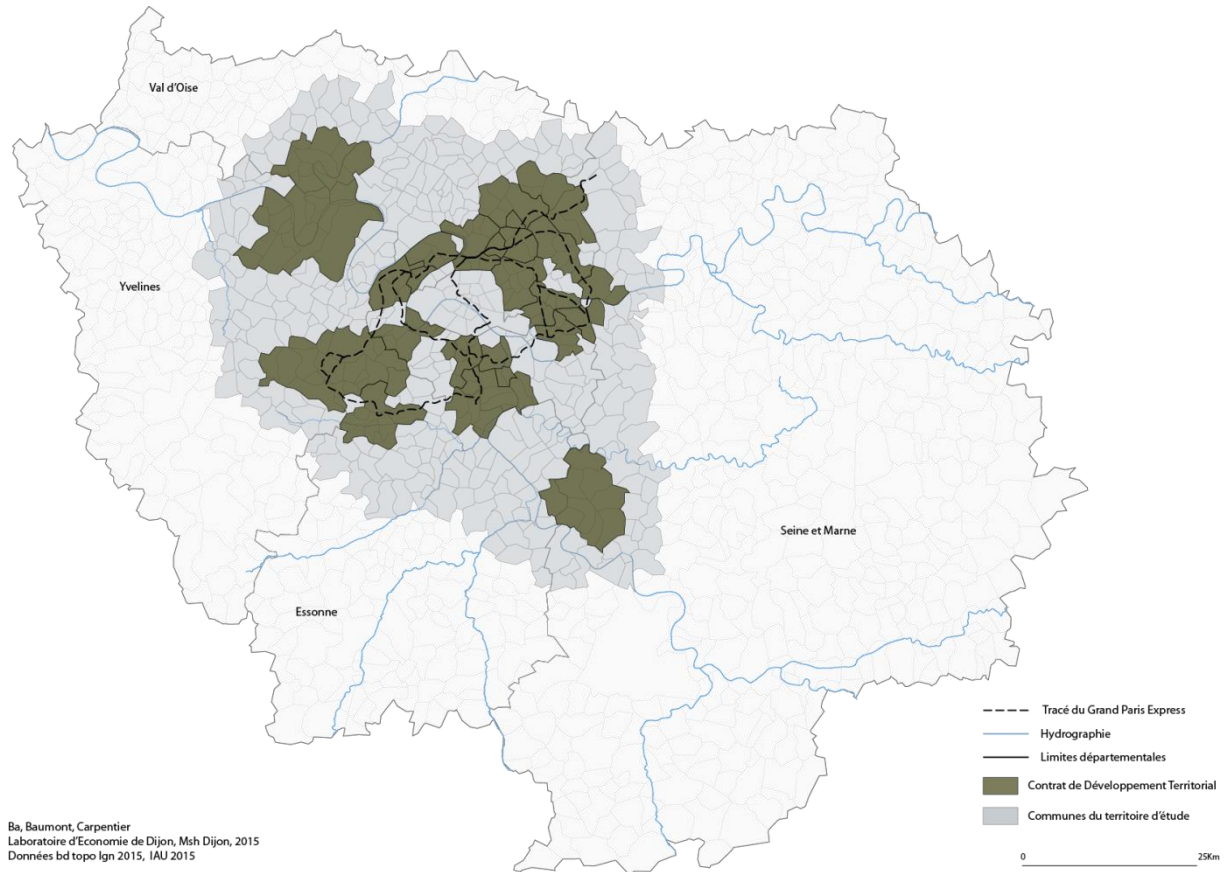
2.3.1 Présentation du périmètre d'étude

La définition du périmètre d'étude constitue la première étape de l'analyse empirique. Plusieurs arguments ont conduit à sa définition. Prendre l'ensemble des communes franciliennes a été exclu, car la localisation des CDT les plus éloignés de Paris reste assez faible (30 km) à l'échelle de la région (*cf.* Carte 2-3) et en tout cas ne couvre pas les parties périphériques des départements de la Grande Couronne. Nous avons également exclu la ville de Paris qui ne fait pas partie, au sens propre, des territoires éligibles, afin de ne pas biaiser les résultats. Compte tenu de sa position centrale et de ses caractéristiques particulières en tant que ville capitale, elle fait figure d'*outlier* par rapport aux autres communes franciliennes. D'autant plus, que certaines variables n'étant pas disponibles à l'échelle des arrondissements, il aurait fallu considérer la ville dans sa totalité. Dès lors, nous avons choisi de retenir un territoire d'un seul tenant, couvrant l'ensemble des communes appartenant à un CDT et l'ensemble des communes qui leur sont géographiquement proches, c'est-à-dire une frange de deux communes autour des CDT. Ce dernier critère délimitant en réalité la partie périphérique de notre aire d'étude tout en ne créant pas une rupture brusque avec les communes des CDT les plus éloignées du cœur de la région.

Au final, notre échantillon est constitué de 483 communes de la région Ile-de-France. L'échantillon comprend 156 communes qui ont décidé d'intégrer un CDT (ce sont les communes qui seront par la suite appelées CDT). Le restant, soit 327 communes, ne font pas partie d'un CDT (ce sont les communes appelées par la suite hors CDT). Ces dernières sont utilisées à double titre : celui de la continuité territoriale (ne pas avoir de vide entre les CDT) et celui des effets de bord (considérer que le périmètre du Grand

Paris ne se termine pas brusquement aux frontières des CDT périphériques. Le périmètre d'étude présenté par la Carte 2-4, couvre une superficie de 3508 km² (soit 29,08 % de la superficie de la région Ile-de-France).

Carte 2-4 : Le périmètre d'étude et le projet « Grand Paris »



La matrice d'interactions ou de pondérations spatiales W , est de dimension 483×483 . Généralement, la matrice de pondération est construite à partir de critères géographiques (contiguïté, distances seuils, inverse de la distance...). Dans notre étude, nous retenons cette dimension mais nous considérons également une caractérisation institutionnelle pour tenir compte de la structuration de ce territoire.

La dimension géographique est importante pour tenir compte des effets de débordement. Pour chaque commune i , nous considérons que son voisinage est composé

de l'ensemble des communes j situées dans un rayon de 5 km. Les éléments w_{ij} de la matrice W sont définis de la manière suivante :

$$w_{ij} = \begin{cases} 1, & \text{si } dist_{ij} \leq 5km \\ 0, & \text{sinon} \end{cases}$$

$dist_{ij}$ étant la distance séparant la commune i et la commune j .

Le choix de la distance seuil de 5 km repose sur les arguments suivants. Avec ce seuil, chaque commune a en moyenne 10 communes dans son voisinage ce qui est un nombre permettant de capter une certaine variabilité. Avec cette matrice, la composition des voisinages retrace l'hétérogénéité des communes selon leur statut : CDT ou hors CDT (cf. Tableau 2-1). Une commune CDT compte en moyenne 7 communes voisines CDT et 4 communes voisines hors CDT. Pour les communes hors CDT, elles ont en moyenne 8 communes voisines hors CDT et 2 communes voisines CDT.

La dimension institutionnelle rend compte d'autres facteurs susceptibles d'introduire de la dépendance entre les comportements des communes vis-à-vis de leur participation, ou non, au Grand Paris. Il s'agit notamment ici de l'expérience et de la capacité des communes à s'engager dans des projets communs. Les Etablissements Publics de Coopération Intercommunale (EPCI) sont, en effet, des regroupements de communes ayant pour objet l'élaboration de « *projets communs de développement au sein de périmètres de solidarité* ». L'appartenance à un EPCI définit des compétences intercommunautaires pour la mise en œuvre de projets sur l'ensemble du périmètre de l'EPCI. L'intercommunalité favorise l'émergence et la faisabilité des projets (Becouze et al., 2013), notamment en termes techniques et financiers. L'appartenance à un EPCI soutient et développe alors un comportement plutôt coopératif – ou facilitateur - entre les communes qui prennent l'habitude de travailler ensemble. Tout nouveau dispositif proposé qui nécessite une réponse territoriale organisée peut alors être plus facilement saisi par les communes déjà engagées dans des projets communs.

Le critère retenu pour définir le voisinage institutionnel est l'appartenance à un EPCI. Ainsi, deux communes sont voisines si elles appartiennent à un même EPCI. Les éléments w_{ij} de la matrice institutionnelle sont définis de la manière suivante :

$$w_{ij} = \begin{cases} 1, & \text{si les communes } i \text{ et } j \in \text{au même EPCI} \\ 0, & \text{sinon} \end{cases}$$

Selon ce critère de voisinage, chaque commune a en moyenne 7 communes voisines et 43 communes du périmètre d'étude n'appartiennent à aucun EPCI. Plus précisément, les communes CDT ont en moyenne 5 communes voisines CDT et une commune voisine hors CDT. Pour les communes hors CDT, elles ont en moyenne 6 communes voisines hors CDT et une commune voisine CDT (cf. Tableau 2-1). La structure du voisinage avec la matrice institutionnelle, montre que dans un EPCI, on retrouve à la fois des communes CDT et des communes hors CDT. C'est aussi une matrice symétrique, car chaque commune appartenant à un EPCI est voisine des autres communes de l'EPCI.

Tableau 2-1: Description des matrices de voisinage

Types de matrices	Types d'interactions	Min	Moyenne	Max
Matrice basée sur la distance au seuil de 5km	CDT-CDT	1	7,02	17
	CDT-hors CDT	0	3,91	11
	Hors CDT-hors CDT	1	7,76	19
	hors CDT-CDT	0	1,87	12
Matrice institutionnelle basée sur les EPCI	CDT-CDT	0	5,3	12
	CDT-hors CDT	0	1,35	13
	Hors CDT-hors CDT	0	6,3	13
	hors CDT-CDT	0	0,64	10

2.3.2 Présentation des variables

Les différentes variables retenues dans l'étude sont décrites ci-après et les principales statistiques descriptives sont reportées dans le Tableau 2-2. Il s'agit de données de l'année 2010, date de constitution du Grand Paris, qui sont extraites de différentes bases de données Insee (Base permanente des équipements : transports

tourisme, Revenus fiscaux localisés des ménages, Bases de données et fichiers détail du recensement de la population 2010), Base de données fiscales de DGFIP¹⁴ et la Datar¹⁵ (pour le classement des ZRR).

- Les déterminants de la décision des communes d'intégrer ou non un CDT sont décrits par les variables suivantes.

- La variable POPULATION donne la taille de la commune. Elle constitue la variable clé pour la planification des services à offrir. En effet, le décideur politique local est amené à tenir compte des attentes des citoyens en matière d'éducation (système scolaire), en bien-être (équipement culturel ou de loisir, cadre de vie,...). On note que les communes CDT sont plus peuplées avec une moyenne de 28 156 habitants contre 11 651 habitants pour les communes hors CDT.

- La variable REVENU représente le revenu médian annuel des ménages dans la commune. C'est un indicateur de la richesse de la commune et qui conditionne les recettes fiscales à travers les taux de taxes votés par les communes. Les communes hors CDT sont plus riches, avec, en moyenne, un revenu médian de 26 063 € contre 22 340 € pour les communes CDT.

- La variable DETTE désigne le niveau d'endettement par habitant de la commune qui comprend les emprunts et dettes à long et moyen termes restant dus au 31 décembre. Cette variable contribue à déterminer le ratio de surendettement de la commune et représente un indicateur de la santé financière de la commune. Le niveau d'endettement par habitant est plus élevé dans les communes CDT que dans les communes hors CDT avec, respectivement, 955 € contre 868 € par habitant. On peut s'attendre à ce qu'une commune fortement endettée examine ce que la participation au Grand Paris risque d'impliquer comme dépenses supplémentaires immédiates, avant de s'intéresser aux bénéfices futurs qui en résulteront.

- La variable INVESTISSEMENT par habitant est le rapport du total des ressources d'investissements, rapportées à la population totale de la commune. Elle indique le dynamisme économique de la commune pour l'amélioration des conditions de vie. Elle constitue également un bon indicateur pour mesurer l'attractivité de la commune par rapport à l'implantation de nouveaux habitants et de nouvelles entreprises. *Toutes choses*

¹⁴ Direction générale des Finances publiques (DGFIP)

¹⁵ Délégation interministérielle à l'aménagement du territoire et à l'attractivité régionale (DATAR)

égales par ailleurs, plus la commune investit, plus elle offre des services susceptibles de satisfaire les besoins de la population et des acteurs économiques. Le niveau d'investissement par habitant est plus important dans les communes CDT que dans les communes hors CDT, avec, respectivement, un niveau d'investissement moyen par habitant de 576 € contre 539 €.

- La variable DISTANCE-PARIS mesure la proximité des communes par rapport à la ville de Paris. En effet, cette proximité constitue un avantage comparatif en termes d'attractivité pour les emplois ou de mobilités des populations tout en traduisant les contreparties en termes de coûts (notamment pour le logement) ... Elle renvoie aussi à la densité des équipements, notamment en infrastructures de transport. Généralement, cette distance à la Ville centre permet de capter de manière globale les grandes tendances d'organisation spatiale et de formes urbaines liées à la ville centre en termes de valeurs foncières, de densités ou encore d'accessibilités... En moyenne, les communes CDT se situent à 18 km de Paris contre 24 km pour les communes hors CDT.

- La variable STATIONS DE TRAINS indique le nombre de stations de RER dans les communes. Au total, le périmètre d'étude compte 175 stations de RER. Les communes CDT sont plus denses en termes de stations RER. On recense 78 stations RER dans les communes CDT ; soit au moins une commune sur deux qui dispose d'une station RER. Pour les communes hors CDT, on dénombre 97 stations RER ; soit une commune sur trois qui dispose au moins d'une station RER.

- La variable FUTURE STATION DU GPE est une variable *dummy* indiquant si la commune va accueillir une nouvelle gare du GPE. On recense au total 54 stations dans le périmètre de l'étude. 80 % de ces futures stations se trouvent dans les communes CDT contre 20 % dans les communes hors CDT. Cette observation est légèrement contre intuitive, car le projet se présente comme un projet de gare GPE, articulant fortement les projets d'aménagement du territoire, c'est-à-dire les CDT, autour des gares GPE selon les termes « Quartiers de Gare du Grand Paris » employés dans le projet¹⁶.

- Les communes peuvent être engagées dans un ensemble de politiques urbaines spécifiques comme les politiques de renouvellement urbain, ciblées sur les quartiers en

¹⁶ La carte des Quartiers de Gare est notamment accessible sur le site de la SGP : http://www.societedugrandparis.fr/wp-content/uploads/2015/05/carte-grand-paris_sgp-apur.pdf

difficultés de type « *Zones Urbaines Sensibles* » pour la revalorisation du tissu social et « *Zones Franches Urbaines* » pour la revalorisation du tissu économique. Les variables *dummy* ZUS et ZFU indiquent l'existence ou non dans la commune, respectivement des quartiers classés ZUS et/ou ZFU. Ces indicatrices complètent les caractéristiques des communes à deux niveaux. Premièrement, elles précisent la situation économique et sociale particulière de la commune (précarité, taux de chômage élevé, décrochage scolaire,...) et les possibilités d'attraction de la commune en termes de créations d'entreprises, à travers les avantages fiscaux accordés dans ces types de zones. Deuxièmement, elles donnent aussi une information sur l'engagement des communes dans des réponses spécifiques à leur territoire notamment en termes de logements sociaux mais aussi de politiques sociales et culturelles – objectifs forts du Grand Paris. Sur l'ensemble des 483 communes de la zone d'étude, 99 communes, soit 20 % des communes, ont des quartiers classés ZUS dont 55 sont hors CDT et 44 sont des communes CDT. Les quartiers ZFU se retrouvent dans 31 communes, soit 6 % des communes, dont 22 communes hors CDT et 9 communes CDT.

- Une autre variable institutionnelle est la couleur politique (MAJORITE MUNICIPALE) du maire de la commune. En 2010, 57 % des communes étaient gérées par des maires de sensibilité de droite (UMP, Divers Droite, Liste Majoritaire,...). Les communes gérées par les maires de sensibilité de gauche (Parti Socialiste, Divers Gauche, Parti Communiste, les Verts,...) représentent 25 % de communes. Enfin, les communes administrées par le « Centre » (Nouveau Centre, MoDem) représentent 5 % des communes. Le reste des communes, soit 13 %, ont des maires « Sans Etiquette ». Pour les communes CDT, 53 % d'entre elles sont administrées par un maire de sensibilité de gauche, contre 36 % par un maire de sensibilité de droite. Ces pourcentages sont respectivement de 25 % et 57 % pour les communes hors CDT.

Tableau 2-2 : Statistiques descriptives des variables explicatives

VARIABLES	Statut	Min	Max	Moyenne	Ecart-type
POPULATION en 2010	Hors CDT	64	104149	11651	830
	CDT	479	112043	28156	1936
	Ensemble	64	112043	16982	912
REVENU MEDIAN (euros) en 2010	Hors CDT	9720	44855	26063	309
	CDT	9538	39453	22340	550
	Ensemble	9538	44855	24861	286
DETTES (euro/habitant) en 2010	Hors CDT	0	11318	867	48
	CDT	0	2905	955	51
	Ensemble	0	11318	896	36
INVESTISSEMENT (euro/habitant) en 2010	Hors CDT	52	6062	538	33
	CDT	89	4224	575	37
	Ensemble	52	6063	551	25

Tableau 2-3 : Statistiques descriptives des variables explicatives (suite)

VARIABLES		Hors CDT	CDT	TOTAL
DISTANCE DE PARIS (KM)	1st Qu.	18,08	10,84	15,16
	médiane	24,26	17,93	22,21
	3rd Qu	30,31	24,93	28,76
STATIONS DE TRAIN (nombre)	0-station	230	78	308
	1-station	77	47	124
	2-stations	15	20	35
	3-stations	5	8	13
	4-stations	0	2	2
	5-stations	0	1	1
FUTURE STATION DU GPE (Dummy)	Non	316	113	429
	Oui	11	43	54
ZUS (Dummy)	Non	283	101	384
	Oui	44	55	99
ZFU (Dummy)	Non	318	134	452
	Oui	9	22	31
MAJORITE MUNICIPALE (Dummy)	Centre	17	12	29
	Droite	187	57	244
	Gauche (référence)	81	82	163
	Sans étiquette	42	5	47

2.4 Résultats des estimations

Pour apprécier la présence des interdépendances stratégiques dans la prise de décision des communes pour intégrer un CDT, nous estimons deux types de modèles : un modèle auto-logistique avec deux matrices de poids différentes et un modèle logistique standard qui sert de comparaison. Dans la spécification du modèle logistique standard, les interdépendances stratégiques potentielles entre les communes sont ignorées, ce qui peut conduire, si elles sont présentes, à des estimations biaisées des paramètres du modèle. Les spécifications auto-logistiques sont estimées avec la méthode du pseudo maximum de vraisemblance.¹⁷

2.4.1 Présentation des résultats

Le tableau 2-4 présente l'ensemble des résultats. Pour les deux estimations auto-logistiques, les résultats montrent que le paramètre d'interdépendance stratégique ρ est significatif et positif dans les deux variantes du modèle : avec un coefficient estimé de 0,538 significatif à 1 % avec la matrice institutionnelle et 0,874 significatif à 5 % avec la matrice géographique. Ces deux résultats conduisent à rejeter l'hypothèse nulle d'absence d'interdépendance entre les communes dans leurs prises de décision d'intégrer un CDT. Le rejet de l'hypothèse nulle permet de vérifier l'hypothèse des interdépendances horizontales selon laquelle la décision d'une commune d'intégrer un CDT dépend des décisions prises par les communes voisines. Ce résultat rend compte de l'impact de la proximité géographique, mais surtout de la soutenabilité des appartenances institutionnelles dans l'engagement des acteurs publics. Le signe positif du paramètre d'interaction ρ indique que les communes appartenant au même voisinage prennent des décisions similaires : il s'agit donc de formes d'interactions que l'on peut qualifier de complémentaires et plutôt associées à des comportements coopératifs. Ainsi, la probabilité qu'une commune intègre un CDT augmente si ses communes voisines ont déjà intégré un CDT.

Ces résultats sont confirmés par le calcul de l'indice de Moran qui permet de tester la présence de l'autocorrélation spatiale de la variable dépendante, et donc de vérifier s'il y a des ressemblances ou des similitudes dans les décisions de participation des

¹⁷ Les calculs ont été faits avec le logiciel R

communes. Les résultats montrent que les indices de Moran obtenus avec les deux matrices sont positifs et largement significatifs avec un indice égal à 0,655 avec la matrice institutionnelle et 0,479 avec la matrice géographique. Ce résultat signifie que les décisions adoptées par une commune et ses voisins géographiques ou institutionnels sont similaires. En plus, en comparant les indices, cette similarité est plus importante avec la matrice institutionnelle.

Pour les variables explicatives, les résultats montrent que certaines variables ont des effets significatifs sur la décision des communes, mais avec des implications différentes. Nous présentons dans les paragraphes suivants, les interprétations des coefficients des modèles auto-logistiques et les effets marginaux qui en découlent.

Concernant l'effet de la proximité de la commune, la DISTANCE par rapport à Paris est significative et négative. Cela signifie que, plus la commune est proche de Paris, plus sa probabilité d'intégrer un CDT est grande. L'influence du projet Grand Paris, même si la ville de Paris est exclue des actions d'aménagement, est bien un projet intégrateur de la ville de Paris et des communes de la petite couronne les plus proches.

Les estimations du coefficient associé à la variable POPULATION sont significatives avec les deux spécifications auto-logistiques. Ce résultat implique que plus les communes sont importantes en termes de taille, plus elles sont intéressées par les CDT. Ainsi avec une augmentation de 1 % de la population, les probabilités d'intégration d'une commune dans un CDT augmentent de 3,5 % avec la matrice institutionnelle et 6,2 % avec la matrice géographique. La prise en compte des interdépendances permet ainsi de mettre davantage en exergue l'effet de la taille de la population dans les décisions des décideurs politiques locaux.

La valeur estimée du coefficient associé à la variable DETTE est négative et significative uniquement dans le modèle avec la matrice géographique. Les communes dont le niveau d'endettement par habitant est important sont moins prédisposées à intégrer un CDT. Ce résultat peut s'expliquer par le fait que l'appartenance à un EPCI tend à lisser l'influence des variables budgétaires communales. En considérant les interactions avec la matrice géographique, une hausse de 1 % dans la variable DETTE se traduit par une baisse de 4 % des probabilités d'intégration d'une commune à un CDT.

En ce qui concerne l'effet d'investissements sur la décision des communes, la variable INVESTISSEMENT est faiblement significative mais toujours positive.

Concernant l'accessibilité de la commune, la présence d'une STATION DE TRAIN dans la commune a un effet significatif et positif sur la décision d'une commune d'intégrer un CDT. Avec la prise en compte des interactions, le résultat signifie qu'une présence d'une STATION DE TRAIN supplémentaire dans la commune entraîne une augmentation de ses probabilités d'intégrer un CDT de 0,03 avec la matrice institutionnelle et 0,053 avec la matrice géographique par rapport à une commune sans STATION DE TRAIN.

C'est également les mêmes effets avec la variable FUTURES STATIONS. Les résultats montrent que l'aménagement d'une future station GPE a un effet positif et significatif dans les décisions des communes d'intégrer un CDT. Cela signifie que la présence d'une future station de GPE augmente les probabilités d'intégration de ces communes de 0,14 point avec matrice institutionnelle et 0,24 point avec la matrice géographique par rapport aux communes n'ayant pas de future station.

Enfin, l'estimation du coefficient associé à la MAJORITE MUNICIPALE est significative et positive dans les trois spécifications (la MAJORITE MUNICIPALE « Gauche » est la référence). Les résultats montrent que les communes dirigées par une majorité de gauche sont plus susceptibles d'intégrer un CDT que celles dirigées par la droite. Ce résultat est sans doute à rapprocher de l'attitude des maires de « droite » qui, pendant plusieurs mois, avaient boycotté le syndicat mixte Paris-Métropole¹⁸ contrôlé à l'époque par la gauche. Cependant, la prise en compte des interdépendances relativise l'effet de la couleur politique sur les décisions des décideurs locaux. Ainsi, les communes gérées par un maire de sensibilité de droite voient leurs probabilités d'intégrer un CDT diminuées de 0,067 point avec la matrice institutionnelle et 0,173 point avec la matrice géographique par rapport aux communes dirigées par un maire de sensibilité de gauche.

¹⁸ Après des mois de boycott, Patrick Devedjian et Valérie Pécresse annoncent qu'ils laisseront les collectivités franciliennes dominées par la droite adhérer au syndicat d'études Paris Métropole, pour l'heure contrôlé par la gauche, http://www.lexpress.fr/actualite/politique/la-droite-rejoint-enfin-paris-metropole_900391.html

En revanche, nous ne trouvons pas d'influence significative des variables associées à la présence des politiques urbaines (ZUS et ZFU) et au revenu médian. Ces variables n'ont aucun effet sur la décision des communes d'intégrer un CDT, quel que soit le modèle utilisé.

Le résultat des estimations dans son ensemble montre qu'à chaque fois que les interdépendances sont considérées, l'effet des variables explicatives est important et ce qui vient renforcer l'idée selon laquelle il existe des interactions dans les prises de décision des décideurs politiques locaux.

Par ailleurs, le rapport des log de vraisemblance entre le modèle réduit (logistique) et le modèle complet (auto-logistique), sous l'hypothèse que les paramètres des deux modèles sont conjointement nuls, à l'exception de WY, est largement significatif, plaidant ainsi en faveur du modèle auto-logistique.

Par rapport à la capacité prédictive du modèle, les résultats présentés dans le Tableau 2-5 montrent que la spécification auto-logistique a une meilleure capacité prédictive ou score d'exactitude, avec un score de 87 % contre 78 % avec le modèle logistique. Le taux de vrais positifs est de 83 %, avec le modèle auto-logistique, contre 75 %, avec le modèle logistique, soit une amélioration de 8 %. Le taux de faux positifs passe de 21 %, avec le modèle logistique, à 11 %, avec le modèle auto-logistique, soit une réduction de 10 points.

Tableau 2-4 : Résultats des estimations logistique standard et autologistic

	Modèle logistique standard		Modèle autologistic			
	Coef. estimé	Effet marginal	Matrice institutionnelle		Matrice géographique	
			Coef. estimé	Effet marginal	Coef. estimé	Effet marginal
INTERCEPT	2,102 (6,647)	-	5,832 (8,821)	-	0,677 (6,792)	-
DISTANCE DE PARIS (KM)	-0,031* (0,017)	-0,006*	-0,040* (0,024)	-0,003	-0,022 (0,017)	-0,004
LOG(POPULATION)	0,343** (0,140)	0,062**	0,442** (0,181)	0,035**	0,350* (0,142)	0,062**
LOG(REVENU MEDIAN,2010)	-0,594 (0,616)	-0,108	-1,267 (0,827)	-0,102	-0,490 (0,626)	-0,087
LOG (DETTE, 2010)	-0,210** (0,087)	-0,038**	-0,138 (0,117)	-0,011	-0,224* (0,088)	-0,04**
LOG (INVESTISSEMENT, 2010)	0,366** (0,183)	0,066	0,402* (0,234)	0,032	0,357. (0,185)	0,064
GARE (STATIONS DE TRAIN)	0,302* (0,154)	0,055*	0,380* (0,210)	0,03*	0,298. (0,156)	0,053*
FUTURE GARE GPE (RÉF. NO STATION)	1,354*** (0,405)	0,246***	1,746*** (0,478)	0,14***	1,321** (0,410)	0,236***
ZUS (REF. NO ZUS)	-0,174 (0,363)	-0,032	0,110 (0,460)	0,009	-0,182 (0,367)	-0,032
ZFU (REF. NO ZFU)	0,531 (0,546)	0,096	0,523 (0,673)	0,042	0,525 (0,551)	0,094
NUANCE POLITIQUE : CENTRE (REF. GAUCHE)	-0,494 (0,474)	-0,09	-0,849 (0,611)	-0,068	-0,480 (0,478)	-0,086
NUANCE POLITIQUE : DROITE (REF. GAUCHE)	-0,933*** (0,281)	-0,169***	-0,833** (0,375)	-0,067**	-0,971*** (0,283)	-0,173***

NUANCE POLITIQUE : SANS_ETIQUETTE (REF. GAUCHE)	-0,990* (0,587)	-0,18	-0,723 (0,847)	-0,058	-0,987* (0,586)	-0,176
PARAMETRE SPATIAL			0,532*** (0,053)		0,874* (0,399)	
INDICE DE MORAN				0,655***		0.479***
AIC		495,580		309,34		492,829
LOG -LIKELIHOOD		-234,7883		-140,6724		-232,415
signif codes : *** : 1 % ; ** :5 % ; * :10 %; Std.Error entre ()						

Tableau 2-5 : Comparaison des capacités prédictives des modèles

		Estimation avec le modèle logistique		Estimation avec le modèle auto-logistique			
				Matrice institutionnelle		Matrice distance	
Situation observée		Prédiction		Prédiction		Prédiction	
		hors CDT	CDT	hors CDT	CDT	hors CDT	CDT
hors CDT		302	25	303	24	282	45
CDT		80	76	35	121	98	58

Conclusion

Dans ce chapitre nous nous sommes intéressés à la problématique des effets d'interaction dans le processus de sélection des bénéficiaires d'une politique publique en général. Plus précisément, nous avons étudié les effets d'interaction dans les décisions de participation des communes dans les Contrats de Développement Territorial (CDT) du projet Grand Paris.

Partant d'un modèle théorique développé par (Brueckner, 2003), nous avons proposé un modèle auto-logistique afin de tester la présence de ces effets dans les décisions de participation des communes à un CDT. Le modèle empirique développé est estimé avec deux types de matrices : une matrice géographique et une matrice institutionnelle. Les résultats montrent dans les deux cas, que les paramètres d'autocorrélation spatiale permettant de mesurer les effets d'interaction sont significatifs et positifs. Ces résultats confirment ainsi l'existence des effets d'interaction dans le processus de constitution des CDT et signifient que les communes voisines prennent des décisions similaires.

De manière générale, ces résultats vérifient également l'hypothèse d'existence d'effet d'interaction dans le processus de sélection des bénéficiaires dans le cadre d'une politique non expérimentale que nous avons développées dans le Chapitre 1.

Ce chapitre pose les jalons pour une identification des logiques politico-économiques qui sous-tendent la participation des acteurs dans les CDT compte tenu des enjeux et de l'organisation institutionnelle du territoire. Il peut contribuer à alimenter les recherches sur le Grand Paris, en plus des nombreux travaux menés, notamment dans le cadre du Conseil Scientifique, pour évaluer le potentiel économique du projet dans les années à venir en questionnant la rentabilité de l'infrastructure du Grand Paris Express. À notre connaissance, il s'agit du premier document s'intéressant au processus de constitution des Contrats de Développement Territorial.

Une des limites majeures de ce chapitre réside dans le caractère non généralisable du modèle empirique du fait que le choix de la spécification auto-logistique a été guidé par le processus de constitution des CDT. Cependant, on peut l'adapter avec spécification de type probit ou logit spatial en fonction de la nature des données disponibles.

Les analyses menées dans ce chapitre pourront être renforcées par des données qualitatives, notamment collectés auprès des principaux acteurs, que ce soit les maires, les

responsables de la société du Grand Paris, ou bien encore les préfets afin d'affiner la logique socio-économico-politique qui sous-tend la participation des communes aux CDT.

Le projet Grand Paris est un vaste champ de recherche, outre la constitution des CDT, nous pouvons nous interroger sur d'autres aspects du Grand Paris tel que ses effets sur l'économie, le logement, la planification urbaine ... Nous proposons dans le chapitre suivant, une évaluation des effets *ex ante* des CDT sur les dépenses publiques d'équipement des communes.

Chapitre 3 La prise en compte des interactions dans l'évaluation *ex ante* : un modèle de doubles différences spatial appliqué au Projet du Grand Paris

Introduction

Les effets *ex ante* sont les effets induits par une politique bien avant sa mise en œuvre. Ils correspondent à ce qu'on appelle communément « effets d'annonce », « effets d'aubaine » ou « effets d'anticipation » et apparaissent, généralement, entre l'annonce de la politique et sa mise en application. Ce décalage temporel est déterminant dans une politique, car c'est la période que les bénéficiaires potentiels peuvent mettre à profit pour développer des comportements d'anticipation ou d'adaptation. En fonction de leurs caractéristiques, celles de l'environnement dans lequel elles vivent et leurs niveaux d'information sur la politique, les cibles potentielles adaptent leurs comportements, soit pour profiter au maximum des avantages attendus de la politique, soit pour contourner ou atténuer les effets négatifs qui seront induits par la politique.

C'est également au cours de cette période que les décideurs interagissent sur leurs politiques avec les institutions politiques (gouvernement, parlement ...), les conseils locaux (conseil régional, municipal...) pour discussions, amendements et votes. C'est aussi l'occasion, à travers des séances d'informations, de réunions de sensibilisation, d'échanger et de partager avec les différentes parties prenantes telles que les citoyens, les acteurs économiques, les organisations de la Société Civile, les syndicats, le patronat, les medias... sur les enjeux, la portée, les bénéfices et les coûts de la politique. Ainsi, les décideurs peuvent mettre à profit cette période pour procéder à une évaluation *ex ante*.

L'évaluation des effets *ex ante* est relativement assez peu traitée dans la littérature, comparativement à l'évaluation des effets *ex post*. L'évaluation *ex ante* est une évaluation prédictive dont l'avantage est de permettre aux décideurs, d'une part, de tirer les premiers enseignements sur la politique, particulièrement sur son acceptation sociale, sa portée économique, sociale et, d'autre part, de procéder, au besoin, à des ajustements sur les objectifs, les cibles, la durée, mais surtout éviter de lancer des politiques coûteuses et inefficaces (Todd et Wolpin, 2010). On pourrait également avancer que l'évaluation de l'effet global d'une politique publique ne doit pas se limiter uniquement aux effets *ex post*. Elle doit tenir compte des effets induits en situation *ex ante*.

Sur le plan méthodologique, l'évaluation *ex ante* « s'appuie sur une théorie explicative des comportements. Ceci est essentiel pour analyser dans quelle mesure les résultats d'une politique, dans un contexte donné, peuvent être généralisés » (Gignoux, 2006, p. 61). Selon Dinardo et Lee (2011), les effets *ex ante* ne nécessitent pas une expérience pour qu'ils se

produisent. Ces auteurs soulignent par ailleurs, que contrairement à l'évaluation *ex post*, la crédibilité d'une évaluation *ex ante* dépend de la crédibilité du modèle statistique sur le comportement des individus et de l'environnement auquel ils sont soumis.

Des études empiriques portant sur l'évaluation *ex ante* ont été menées sur différentes politiques (travail, santé, fiscal,...). L'objectif poursuivi est de révéler si effectivement des comportements d'anticipation se sont produits et/ou des effets d'anticipation ont pu être mesurés. Blundell et *al.* (2011) analysent les effets d'annonce de la mise en œuvre de la réforme « *Working Families' Tax Credit –WFTC* », sur la participation des mères monoparentales au marché de travail, au Royaume Uni. A travers une estimation par la méthode *Difference-in-Difference*, ils montrent qu'il existe des effets d'annonce positifs sur les décisions d'emploi éligibles (16 heures de travail ou plus par semaine). Ils soulignent que l'évaluation des effets de la réforme par la méthode avant et après, qui ignorerait ces effets d'annonce produira des estimations biaisées (sous estimées passant de 15 à 35 pour cent).

Crépon et *al.* (2010) s'intéressent aux effets d'anticipation des demandeurs d'emploi dans la ville de Paris suite à une notification d'assister à une formation de la part de l'ex ANPE (Pôle Emploi). Ils soulignent que la notification peut avoir des incidences sur les stratégies de recherche d'emploi du demandeur d'emploi. Ce dernier peut cesser de chercher du travail pour attendre la formation, ou bien multiplier ses recherches, ou encore accepter tout emploi pour ne pas participer à la formation. Les auteurs utilisent les données sur les périodes de chômage issues du Fichier Historique Statistique entre 2003/2004 et 2008 et s'appuient sur un modèle appelé "*timing of Event*" qui exploite la variation du traitement dans le temps par rapport au résultat, car dans ce cas précis la notification est individuelle et dynamique. Les résultats montrent que la notification a un effet positif sur la formation et un effet négatif sur la sortie du chômage.

Dans le cadre de l'évaluation de la réforme « *Finnish Corporate Income Tax Reform of 2005* » portant sur une hausse des impôts sur les dividendes, Seppo et *al.* (2008) étudient comment les entreprises ayant leur siège en Finlande ont modifié leurs comportements par rapport à la redistribution des dividendes, aux investissements et au financement de la dette en prévision de la réforme fiscale de 2005, en 2003 et 2004. Ils estiment l'effet d'anticipation des entreprises à travers un modèle *Difference-in-Difference* avec des données couvrant la période 1999-2004. Les résultats montrent que les entreprises ont augmenté leurs versements de dividendes de manière significative avant la réforme.

Alpert (2012), dans le cadre d'une réforme médicale aux USA, estime les effets d'anticipation du passage du programme « *Medicare Part D* » sur l'utilisation des médicaments sur ordonnance, annoncé en Décembre 2003, et mise en œuvre en 2006. Il utilise deux sources de données (*Medicare Current Beneficiary Survey Cost* pour la période 2001-2006 et *the Medical Expenditure Panel Survey* pour la période 1997-2007) et mobilise la méthode *Difference-in-Difference* pour estimer l'effet d'anticipation. Il trouve que l'effet d'annonce est négatif sur l'utilisation des médicaments sur ordonnance et que sa prise en compte réduit l'effet total estimé du programme au moins de la moitié. Par conséquent, ignorer les effets d'anticipation à l'annonce du programme *part D* en 2003 surestime son impact dans l'utilisation des médicaments.

Bergemann et al. (2011) étudient le comportement des travailleurs en Allemagne selon leur origine avant leur intégration au programme « *Active Labor Market Policy Programs* ». Ils montrent à travers un modèle de *propensity score matching*, que l'effet *ex ante* est positif sur les efforts de recherche d'emploi et négatif sur le salaire minimum que le demandeur d'emploi est prêt à accepter. Ils trouvent que le salaire minimum accepté par les allemands baisse de 3 % et pour les étrangers (Européens de l'Est et les Russes), cette baisse est de 3,9 %. Ils soulignent, par ailleurs que les efforts de recherche d'emploi sont plus importants pour les étrangers avec une qualification élevée que les Allemands de même niveau, avec respectivement 9 % contre 4,3 %.

Comme en *ex post*, la question des effets d'interaction reste également valable en évaluation en *ex ante*. Nous avons montré dans le chapitre 1 de la thèse, que les effets de voisinage sont importants au sein des groupes sociaux, entre les territoires et que les relations entre les individus sont complexes : elles sont liées à leur propension à interagir, ne sont pas nécessairement symétriques, peuvent être hiérarchiques avec des effets de leader... La politique est elle-même un ensemble d'actions localisées, destinées aux individus. Les interactions qu'ils entretiennent entre eux peuvent influencer leur comportement vis-à-vis de cette politique. Dès lors, lorsque ces effets existent et ne sont pas traités, les estimateurs des paramètres d'intérêts sont susceptibles d'être affectés par trois problèmes économétriques (variables omises, simultanéité, endogénéité). *Ceteris paribus*, ne pas tenir compte de ces interactions positives (resp. négatives) revient à sous-estimer (resp. sur-estimer) les effets d'un programme (Affuso, 2010).

L'évaluation *ex ante* de la politique n'échappe pas *a priori* à ces effets et on peut souligner que la connaissance de la mise en place future d'une action publique peut tout à fait susciter et faire l'objet de nombreuses interactions entre les individus. Ainsi, considérer les effets d'externalités de la politique sur les autres individus, permet de mieux ajuster les effets réels qu'elle induit, bien avant sa mise en œuvre.

Dans le cadre des politiques territorialisées, les communes appartenant à un même territoire présentent souvent des caractéristiques structurelles similaires (structure démographique de la population, caractéristiques culturelles et environnementales...). La proximité entre les communes, qu'elle soit de nature géographique, économique, sociale, politique ..., est à l'origine d'interdépendances entre leurs décisions. Cette proximité favorise également la diffusion des effets de la politique sur les communes voisines.

Par conséquent, toute décision prise par une commune peut induire des externalités qui peuvent impacter positivement ou négativement les communes voisines. Par exemple, la construction d'un centre de loisirs, l'entretien des espaces verts ou l'aménagement des pistes cyclables par une commune, profitent à la fois à sa population et à celles de ses voisins proches. Cela peut conduire ainsi les décideurs politiques locaux à adopter des décisions optimales, qui tiennent compte des interdépendances avec les décisions des autres collectivités.

Deux courants traitent de ces interactions stratégiques, selon qu'ils portent sur l'analyse des décisions publiques ou sur les stratégies électorales¹⁹.

- Dans le cadre des dépenses publiques, l'adoption d'une taxe locale attractive pour l'implantation des entreprises, par une commune, peut affecter la structure économique des communes voisines, soit positivement, à travers la création de nouveaux emplois, soit négativement, à travers la délocalisation des entreprises des communes voisines, ce qui conduit donc à la réduction de leurs assiettes fiscales. Dans ce cas, la décision optimale des communes peut se porter sur des variables stratégiques : le niveau de dépenses publiques, ou leurs niveaux de taxes.

- L'adaptation des comportements des communes peut être également motivée par des enjeux électoraux. En effet, les électeurs d'une commune peuvent comparer les politiques menées par leurs élus à celles menées dans les communes voisines. Ils peuvent se servir des

¹⁹ Les comportements stratégiques résultant des interactions stratégiques entre les collectivités ont été traités sur le plan théorique dans de très nombreux travaux dont on peut trouver une synthèse dans Schone, (2010).

résultats de ces politiques pour mesurer la qualité de leurs élus et, par conséquent, les sanctionner lors des élections. Le comportement des électeurs peut pousser ainsi, les élus d'une commune, à adopter des comportements similaires à ceux de leurs homologues pour être réélus (Schone, 2010 ; Schone et al., 2013).

Sur le plan empirique, ces interactions stratégiques ont fait l'objet de nombreuses applications notamment sur : (i) les interdépendances au niveau des dépenses publiques (Case et al., 1993) entre Etats aux USA, (Schaltegger et Küttel, 2002) entre les cantons en Suisse, (Solé-Ollé, 2006) entre les municipalités en Espagne. (ii) Les interdépendances au niveau de taxation (Redoano, 2003) entre les Etats Européens, (Allers et Elhorst, 2005) entre les municipalités aux Pays Bas, (Feld et al., 2002), entre les régions en France, (Leprince et al., 2005; Freret, 2008), entre les départements en France et (Jayet et al., 2002; Schone, 2010; Schone, 2014; Schone et al., 2013) entre les collectivités locales en France.

L'analyse des interactions entre collectivités peut également porter sur les opérations d'aménagement et/ou sur les incidences fiscales et budgétaires de ce projet. C'est l'objet de la partie empirique de ce chapitre. Un projet d'aménagement d'envergure comme l'est celui du Grand Paris, a été annoncé en 2007. Il vise à doter le cœur de la région Ile-de-France d'une infrastructure de transport urbain – le Grand Paris Express – et s'appuie sur des partenariats locaux avec les communes qui s'engageraient aux côtés de l'Etat et de la Région Ile-de-France pour développer des projets d'aménagement touchant l'habitat, l'emploi, les domaines culturels et environnementaux. Ce projet est à même de transformer de manière durable la région du fait d'enjeux économiques et financiers importants : 70 000 logements neufs par an, 15 000 à 20 000 emplois par an²⁰, 26 milliards d'euros d'investissement en 15 ans. Il ne peut qu'interpeller les communes concernées et impactées : susciter de la concurrence, de la compétition et engendrer des interdépendances stratégiques très fortes entre les communes, entraînant ainsi des comportements d'anticipation des décideurs publics locaux sur plusieurs aspects. Avec les retombées économiques, démographiques et politiques attendues, les communes franciliennes peuvent mettre en place, en *ex ante* des dispositifs attractifs par l'accroissement des investissements publics pour améliorer les services (parking, le transport, crèches,...) et construire des logements. En même temps, le coût de ces actions et le coût dans

²⁰ <http://www.grand-paris.jll.fr/id/121>

la participation au projet (cf. Chapitre 2) peut amener les communes à s'interroger sur leurs engagements.

L'objectif de ce chapitre est de proposer un cadre méthodologique pour réaliser une évaluation *ex-ante* du dispositif de Grand Paris, en prenant en compte les potentielles interdépendances entre les communes. Plus précisément, l'objectif est de voir si le projet Grand Paris a déjà induit des effets d'anticipation sur le niveau de dépenses d'équipement des communes tout en contrôlant les effets d'interdépendances de ces dernières. Pour prendre en compte ces effets dans l'estimation de l'effet du Grand Paris, nous supposons que le niveau des dépenses d'équipement d'une commune est influencé par les dépenses d'équipement de ses voisines. Les effets de ces investissements ne se limitent pas à ses frontières mais ils impactent les populations des autres communes voisines, à travers différents mécanismes de diffusion.

L'enjeu opérationnel de cette étude est de disposer, à la date de 2012, d'une situation de référence qui pourra être confrontée, plus tard, aux effets de la politique du Grand Paris. C'est précisément cet éclairage par les tendances passées qui nous permet, d'une part, d'apprécier en *ex-ante* la pertinence des choix et de projeter, d'autre part, les dynamiques futures.

Pour cela, nous proposons une extension de la méthode *Difference-in-Difference*, avec les méthodes de l'économétrie spatiale. Cette extension permet alors d'estimer l'effet *ex ante* du projet Grand Paris avec la prise en compte des effets d'interdépendances. Nous allons considérer pour chaque type de communes les effets d'interaction avant et après le programme à travers une spécification de type *spatial autoregressive model -SAR*.

La première section sera consacrée à la présentation du programme Grand Paris (3.1). Le modèle théorique de compétition fiscale employé et le modèle empirique associé sont présentés dans la deuxième section (3.2). La troisième section est consacrée à l'identification et l'estimation de l'effet d'anticipation (3.3). Enfin, la quatrième et la cinquième section sont consacrées, respectivement, à la description des données et des matrices de voisinage (3.4) puis aux résultats des différentes estimations économétriques (3.5).

3.1 Présentation du projet du Grand Paris

La région Ile-de-France est l'une des plus grandes régions européennes, avec 4,5 % du PIB de l'Union Européenne, en 2012. Elle se caractérise par son image de dynamisme économique : forte présence d'entreprises multinationales, 7 pôles de compétitivité, dont 4

pôles mondiaux²¹ et son envergure internationale ; la métropole Parisienne étant classée dans le top cinq des métropoles mondiales²² en 2012. Elle reste encore en 2012, la première destination touristique mondiale, avec environ 29 millions de touristes²³.

L'Ile-de-France compte environ 12 millions d'habitants, soit 18,3 % de la population française et elle représente 30 % du PIB national. La population active est estimée à plus de 6 millions d'actifs en 2011 et le taux de chômage moyen est de 8,2 % en 2012²⁴.

Cependant, malgré ses atouts, l'Ile-de-France reste confrontée à un ensemble de difficultés qui pénalisent son dynamisme économique (Carré et Levratto, 2013). Le cadre de vie francilien connaît des limites, les transports rencontrent des difficultés dans leur exploitation – tant au niveau de leur saturation que de leur organisation, la desserte en étoile contribue à augmenter le budget temps destiné aux migrations pendulaires des actifs. Le logement est aussi un point de tensions pour la région, car l'insuffisance d'une offre adaptée contribue naturellement à la hausse des prix des biens immobiliers.

Dans ce contexte, le projet Grand Paris a été initié avec, comme ambition, de permettre à la métropole parisienne de conserver son statut de ville monde, de rester compétitive face à des villes concurrentes comme Londres, New York ou Tokyo, mais aussi à des pôles mondiaux émergents comme Séoul ou Shanghai. L'idée générale du projet est donc d'améliorer le cadre de vie au sens large, afin de permettre à la région capitale de poursuivre sa croissance et de la rendre plus attractive.

Les enjeux et caractéristiques du projet Grand Paris²⁵

Le projet Grand Paris vise à redynamiser la région francilienne, en agissant sur un certain nombre de leviers économiques, afin d'améliorer l'attractivité et la compétitivité de la région parisienne (Prager, 2014). L'action emblématique du projet concerne la construction

²¹ Les 4 pôles mondiaux sont Finance Innovation, Systematic Paris Région, Cap Digital Paris-Région et Medicen Paris Région

²² The Economist : Hot spot : Benchmarking Global City Competitiveness, 2012

: <http://www.managementthinking.eiu.com/sites/default/files/downloads/Hot%20Spots.pdf>

²³ <http://presse.parisinfo.com/etudes-et-chiffres/chiffres-cles/le-tourisme-a-paris-chiffres-cles-2012-edition-2013>

²⁴ Sources : Insee, RP2012 exploitations principales :

<http://www.insee.fr/fr/themes/comparateur.asp?codgeo=reg-11>

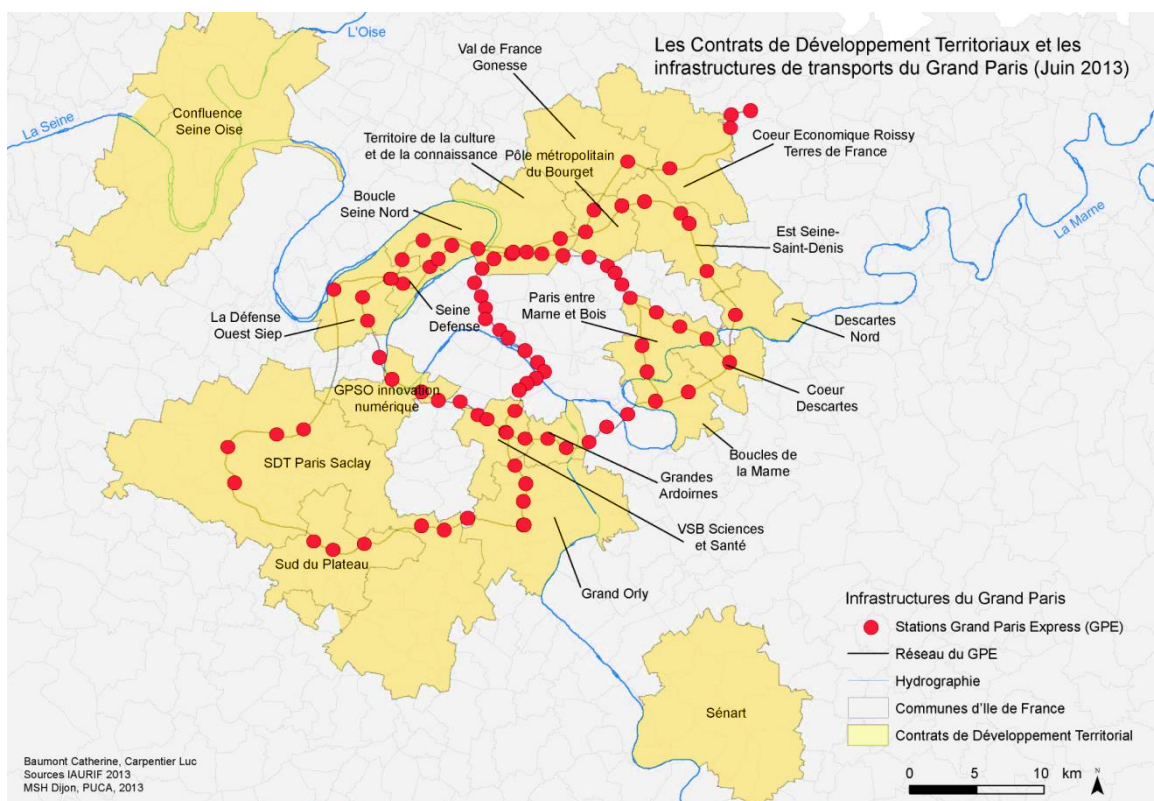
http://www.insee.fr/fr/themes/document.asp?reg_id=20&ref_id=19976

²⁵ Une partie des éléments de cette partie est reprise du chapitre 3.

du métro « Grand Paris Express » (Beaucire et Dauzelle, 2013)²⁶ dont l'objectif est de désenclaver et de faciliter les déplacements entre les territoires situés à la périphérie de la ville de Paris. Cette action s'accompagne d'une stratégie de développement urbain, social et économique des territoires de la région.

La politique des transports du « Grand Paris Express²⁷ » vise à moderniser et à étendre le réseau actuel avec 200 km de lignes installées impliquant 72 gares dont 57 nouvelles gares, la création de 4 nouvelles lignes (15, 16, 17 et 18) (cf. Carte 3-1), et ceci, afin de transporter en moyenne 2 millions de voyageurs par jour. L'enjeu environnemental du GPE est également souligné afin de rendre la substitution voiture-transport collectif plus réaliste dans les territoires franciliens, jusque-là, mal ou peu desservis par les transports urbains.

Carte 3-1 : Le tracé du métro Grand Paris Express et localisation des CDT



²⁶ L'hypothèse, par ailleurs théoriquement fondée est que cette infrastructure de transport va exercer une action structurante pour la région. Ses impacts font notamment l'objet de diverses études dans le cadre du conseil scientifique du Grand Paris. <http://www.societedugrandparis.fr/>

²⁷ Les chiffres donnés dans ce paragraphe viennent de : <http://www.grand-paris.jll.fr/le-projet-du-grand-paris/presentation-generale/id/48>

Sur le plan de l'habitat, le projet vise la construction de 70 000 logements neufs par an, pour faire face à la crise de l'habitat qui sévit dans l'agglomération parisienne, et d'ici 2030, les projections indiquent la construction de 1,2 million nouveaux logements. Ces actions s'inscrivent dans la réduction de la consommation d'énergie avec la construction massive de logements neufs satisfaisant donc les dernières réglementations en matière d'efficacité énergétique.

Sur le plan économique, les projections indiquent que le projet Grand Paris va générer environ 15 000 à 20 000 emplois par an dans le secteur des travaux publics, via la construction et les rénovations du réseau de transport et, sur le long terme, la création de 115 000 à 315 000 emplois à travers le gain d'attractivité de la région. Sur le plan de la Recherche- Développement-Innovation, le projet a pour objectif de faire de la métropole parisienne une vitrine de l'excellence universitaire française et de renforcer l'accessibilité des grands équipements culturels et sportifs. La contribution du projet Grand Paris au PIB français est estimée autour de 140 milliards €.

Cette volonté politique a été matérialisée par la loi n° 2010-597 du 3 juin 2010 relative au Grand Paris à travers l'article 7 portant création de l'établissement public la « Société du Grand Paris » (SGP pour la suite) dont « *la mission principale est de concevoir et d'élaborer le schéma d'ensemble et les projets d'infrastructures composant le réseau de transport public du Grand Paris (...)*»²⁸. Outre la mission de mise en œuvre de l'infrastructure Grand Paris Express, la SGP doit également veiller « *au maillage cohérent du territoire par une offre de transport de surface permettant la desserte des gares du réseau de transport public du Grand Paris*»²⁹. A cela s'ajoute l'article 21 de la même loi portant sur le développement territorial et des projets d'aménagement et le décret d'application n° 2011-724 du 24 juin 2011 relatif au Contrat de Développement Territorial (CDT)³⁰.

Les CDT offrent des possibilités aux communes et aux Etablissements Publics de Coopération Intercommunale (EPCI) concernés par le futur réseau de transport Grand Paris Express, de mettre en place, avec l'Etat, les contrats de développement territorial dont l'objet est de porter des projets de développement stratégique du Grand Paris.

²⁸ <http://legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000022308227&dateTexte=&categorieLien=id>, consultée le 28/09/2015.

²⁹ Article 7 de la loi n° 2010-597 du 3 juin 2010

³⁰ Pour des compléments sur les CDT on pourra se reporter au chapitre 2 de la thèse.

En 2012, le territoire du projet Grand Paris qui exclut la ville de Paris, regroupait 21 CDT, pour un total de 156 communes, soit 13 % des communes de l'Ile-de-France. L'ensemble de ces communes couvre une superficie de 1 196 km², soit 10 % de la superficie de l'Ile-de-France, avec une population de 4 461 000 habitants (38 % de la population régionale) et 2 363 000 emplois (42 % des emplois régionaux).

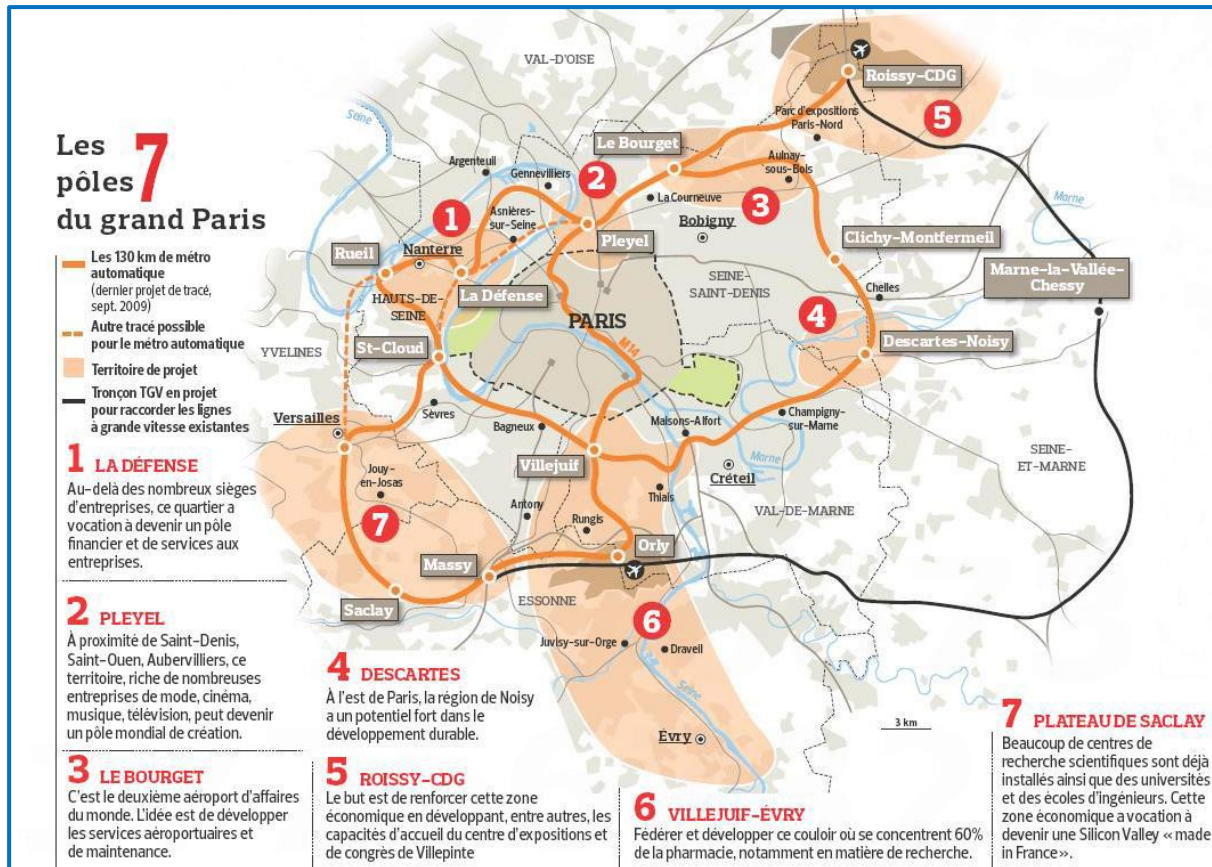
Le CDT est la déclinaison territoriale des objectifs du Grand Paris et définit un nouveau périmètre de réflexion autour des dynamiques métropolitaines avec des objectifs cohérents avec la stratégie globale du Grand Paris³¹. Les différents CDT qui accompagnent les territoires de projet insistent sur trois mécanismes de développement et d'attractivité : (i) améliorer l'accessibilité via les infrastructures de transport, (ii) développer l'innovation en créant un climat de la connaissance et (iii) agir sur le cadre de vie (transports, habitat, environnement et culture).

Par ailleurs, chaque CDT est porteur d'identité d'un territoire et ils sont articulés autour des 7 clusters³² ou pôles de compétitivité thématiques aux alentours de Paris (*cf.* Carte 3-2).

³¹ Flash Eco n°35 : Le Val-de-Marne et ses CDT : un territoire au cœur des projets du Grand Paris, Nov. 2013

³² La description des clusters provient du site <http://www.grand-paris.jll.fr/fr/le-projet-du-grand-paris/-les-clusters/id/50>, consulté le 28/09/2015

Carte 3-2 : Les territoires stratégiques du Grand Paris



Source : <http://www.lefigaro.fr/assets/images/grand-paris.jpg>, page consultée le 14/09/2015

- Le cluster « Plateau de Saclay, pôle de l'Innovation et de la Recherche » axé sur la recherche et l'innovation. Il constitue le plus important campus scientifique et technologique d'Europe. Il regroupe les CDT « Versailles Grand Parc - St-Quentin – Vélizy » et « Paris Saclay Territoire Sud ». Ce dernier compte capitaliser sur la forte présence des établissements scientifiques et universités implantés dans son périmètre pour renforcer et créer le pôle Saclay en tant que campus et pôle de recherche de niveau mondial : le territoire compte regrouper à terme, 25 000 chercheurs, 17 000 étudiants et 160 laboratoires publics et privés.

- Le cluster « Pôle des Echanges internationaux et de l'Événementiel à Roissy-Charles de Gaulle » est naturellement orienté vers les activités soutenues par les échanges internationaux de personnes, touristes, hommes d'affaires, cadres... Il regroupe les CDT « Val de France – Gonesse – Bonneuil », « Cœur économique Roissy Terre de France » et « Grand Territoire Roissy ».

- Le cluster « *Pôle de la Santé* » est un pôle d'excellence dans le domaine des sciences de la vie. Il ambitionne de faire émerger un territoire de référence internationale en matière de recherche et de lutte contre le cancer. Il regroupe trois CDT : (i) le CDT « *Campus Sciences et Santé* », qui va capitaliser sur la présence d'un grand pôle hospitalo-universitaire et d'entreprises du secteur, (ii) le CDT « *Grandes Ardoines* », qui va s'appuyer sur la biotechnologie en relation avec l'ensemble des pôles d'excellence du sud de l'Ile-de-France et (iii) enfin le CDT « *Grand Orly* », qui s'insère dans un territoire riche en projets et en équipements, notamment l'aéroport d'Orly et le marché international de Rungis.

- Le cluster « *La Défense* » est dédié exclusivement à la Finance. Ce pôle représente le premier quartier d'affaire européen et abrite de nombreux sièges sociaux de multinationales (EDF, GDF-Suez, Total, Philip Morris...) et les bureaux des plus grandes banques d'affaires (Société Générale, Axa, HSBC, American Express,...). Il est constitué par deux CDT : le CDT « *Seine Défense* », dont l'un des objectifs phares est « d'accompagner la dynamique générée par le quartier d'affaires » et le CDT « *La Défense Ouest - Les deux Seine* » qui a comme objectif « d'inscrire ce territoire dans une dynamique métropolitaine compétitive, créatrice d'emplois et de répondre aux attentes des populations ».³³

- Le cluster « *le pôle de l'Aéronautique du Bourget* » du Bourget est tourné vers l'industrie aéronautique et le développement du tourisme d'affaires. Il compte capitaliser sur la présence sur l'aéroport d'affaires du Bourget, le Musée de l'Air et de l'Espace, le Parc des Expositions du Bourget. Il est constitué par le CDT « *Pôle d'Excellence Aéronautique* ».

- Le cluster « *pôle de la Création de Saint-Denis Plaine* » est dédié à la création numérique, avec des objectifs de développement des industries culturelles et créatives autour de la *Cité du Cinéma*. Ce cluster correspond au CDT « *Territoire de la Culture et de la Création* ».

- Le cluster « *pôle de la Ville durable autour de la cité Descartes à Marne-la-Vallée* » axé sur le développement durable. Ce pôle va s'appuyer sur le patrimoine naturel du département et sur les équipements issus de la ville nouvelle. Il regroupe les CDT « *Grand Paris Est Noisy-Champ* » et « *Boucles de la Marne* ».

³³ Source : <http://www.grand-paris.jll.fr/id/93>

- En plus de ces territoires stratégiques, le CDT du « Grand Paris Seine Ouest », dont la thématique principale est liée à l'innovation numérique, abrite de nombreux sièges de grandes entreprises nationales et internationales (Thalès, Dassault, Microsoft, ...).

Cette organisation territoriale structurée par les CDT et les pôles stratégiques, complétée par l'offre nouvelle d'infrastructures, constitue le socle du développement économique qui serait impulsé par le projet Grand Paris. On note alors à propos de cette organisation qu'elle s'appuierait sur des complémentarités entre les CDT et les territoires voisins, tout en veillant à ce que les retombées économiques et sociales (emplois, logements,...) soient bénéfiques aux autres territoires voisins hors CDT.³⁴ Ces territoires stratégiques au cœur de la stratégie du Grand Paris sont assez tôt identifiés et connus des acteurs. Ils correspondent, en effet, aux sites de compétitivité déjà connus et sur lesquelles des dynamiques économiques se développent. Le renforcement de ces pôles apparaît alors assez logique.

Ces hypothèses fonctionnelles, assorties des transformations durables que le projet va initier, sont en tout cas susceptibles de créer des réactions de la part des acteurs du territoire.

3.2 Les interactions entre les collectivités : cadre théorique et empirique

3.2.1. La concurrence entre collectivités : éléments théoriques

Dans le cadre du Grand Paris, on suppose que les communes sont en compétition pour la mise en place de dispositifs attractifs sur leur territoire. Cette compétition entre les communes peut être analysée sous l'angle de la concurrence fiscale. Cette concurrence peut générer deux types d'externalités, à savoir les externalités verticales et les externalités horizontales. Ces externalités résultent des effets d'influence de certains acteurs sur les autres. La concurrence verticale s'exerce entre collectivités de niveaux différents alors que la concurrence horizontale s'exerce entre acteurs d'un même niveau et dans ce cas l'externalité horizontale survient lorsque la décision d'une collectivité impacte le budget d'une collectivité

³⁴ Résolution de Paris Métropole sur les Contrats de Développement Territorial, du 5 juillet 2012.

de même rang. Ce sont les externalités horizontales qui nous intéressent ici avec les décisions des communes.

Pour prendre en compte les externalités horizontales entre communes sur des variables stratégiques comme les niveaux de taxes ou les dépenses publiques, nous considérons à l'instar de Jayet et al. (2002), Madiès et al. (2005), le modèle de concurrence fiscale développé par Wildasin (1988)

Soit y un bien privé qui sert de numéraire. Il est produit dans chaque commune à partir du facteur capital (K_i) supposé homogène, parfaitement divisible et mobile. Chaque commune i est caractérisée par une fonction de production $f_i(K_i)$. On suppose que la fonction de production $f_i(K_i)$ est quasi-concave et deux fois différentiables, c'est-à-dire :

$$f_i' > 0 > f_i''$$

Par ailleurs, le facteur capital est rémunéré à sa productivité marginale, notée $f'(K_i)$. Le rendement local pour les facteurs fixes est donné par la relation suivante :

$$f_i(K_i) - K_i f'(K_i) \tag{3-1}$$

Le bien privé produit peut être soit consommé, soit utilisé comme intrant dans la fourniture d'un bien public noté z .

Pour une commune i , la quantité de bien public z_i , peut être mesurée par les dépenses publiques (Wildasin, 1988, p.231). Les dépenses publiques de la commune sont financées à travers une taxe par unité de capital à un taux t_i .

La commune est soumise à une contrainte budgétaire qui impose que le bien public est financé par le produit de la taxe (équation (3-2)) :

$$t_i \cdot K_i = z_i, \forall i = 1, 2 \dots n \tag{3-2}$$

Le capital est rémunéré à sa productivité marginale et il est supposé mobile, de telle sorte que le rendement net par unité de capital noté ρ , est égal à :

$$\rho = f'(K_i) - t_i \forall i = 1, 2 \dots n; \tag{3-3}$$

En situation d'équilibre le rendement net du capital est fonction du vecteur du taux de taxes $\tau = (t_1, \dots, t_n)$.

Soit $\xi = (z_1, \dots, z_n)$ le vecteur des dépenses publiques et $\bar{K} = \sum_i K_i$ le stock de capital, supposé fixe.

On suppose qu'une commune est habitée par des ménages représentatifs en termes de préférences et de revenus. La consommation du bien privé par un ménage dans la commune i notée y_i est égale au rendement local plus le revenu net du capital reçu.

$$y_i = f_i(K_i) - K_i f'(K_i) + \theta_i \rho \bar{K} ; \forall i = 1, 2 \dots n \quad (3-4)$$

Où $\rho \bar{K}$ est le revenu total du capital, θ_i la part du stock du capital détenu par le ménage résident dans la commune i .

Par ailleurs, chaque ménage est caractérisé par une fonction d'utilité $u_i(y_i, z_i)$ décrivant ses préférences en biens privés (y_i) et publics (z_i). Le bien privé y_{it} peut être exprimé en fonction du vecteur (τ) à travers $K_i(\tau)$ et de $\rho(\tau)$.

A l'équilibre de Nash, chaque commune choisit un niveau de dépense publique z_i qui maximise son utilité, sous contrainte des niveaux de dépenses publiques z_j des autres communes $j \neq i$. En d'autres termes, chaque commune choisit un niveau de dépenses optimales en tenant compte de celles des autres communes.

On suppose que la contrainte budgétaire de la commune équation (3-2) peut être satisfaite en situation d'équilibre.

En remplaçant dans l'équation (3-2), K_i par $K_i(\tau)$ i.e. $z_i = z_i(\tau) = t_i K_i(\tau)$ alors le vecteur des dépenses publiques $\xi = (z_1, \dots, z_n)$ est fonction de τ .

Le vecteur ξ^* est un équilibre de Nash pour les dépenses publiques si, pour tout i , $z_i^* \in \xi^*$ est une solution du programme suivant :

$$\text{Max } u_i(y_i, z_i)$$

Sous contraintes :

$$(1) y_i = f_i(K_i[\tau(\xi)]) - K_i(\tau[\xi])f'(K_i[\tau(\xi)]) + \theta_i \rho(\tau[\xi])\bar{K} \quad (3-5)$$

$$(2) \tau(\xi) \text{ est la solution de l'équation (3 - 2)}$$

$$(3) z_j = z_j^*, j \neq i$$

En situation d'équilibre, ce sont les niveaux de dépenses qui sont optimisés, avec des taux d'imposition répondant passivement pour maintenir les budgets des communes en équilibre (Wildasin, 1988, p.234).

En d'autres termes, chaque commune i choisit un niveau de dépenses publiques z_i qui maximise sa fonction de production (équation (3-2)). Le niveau de dépense optimal z_i^* conduit à une solution de l'équation (3-6), qui dépend des niveaux de dépenses des autres communes voisines z_j et de ses propres caractéristiques x_i .

$$z_i = R(z_{-i}, x_i) \quad (3-6)$$

La fonction R représente la fonction de réaction qui permet à la commune i d'avoir les meilleures réponses sur les niveaux de dépenses z_{-i} dans toutes les autres communes. Lorsqu'il y a interaction, la pente de l'équation de réaction (3-6) est non nulle. Ainsi pour tester les interactions, on procède à un test de significativité du coefficient de la pente sous l'hypothèse nulle d'absence d'interaction. La spécification économétrique permet d'estimer empiriquement la pente de la fonction de réaction.

3.2.2. Spécification économétrique

Le modèle formalise les interactions entre les collectivités : les dépenses publiques d'une commune i , notée y_i ³⁵, dépendant de ses propres caractéristiques x_i et du niveau de dépenses des communes voisines y_j .

³⁵ Pour conserver les notations employées dans les modèles d'évaluation des effets de traitement, nous notons y_i le résultat correspondant au niveau de dépenses d'équipement z_i du modèle de concurrence fiscale.

Le modèle empirique proposé spécifie que les dépenses publiques d'une commune i dépendent des niveaux de dépenses publiques des autres communes j et sont exprimées sous la forme linéaire suivante (Brueckner et Saavedra, 2001; Case et *al.*, 1993) :

$$y_i = x'_i \beta + \rho \sum_{j \neq i} w_{ij} y_j + u_i \quad (3-7)$$

Il s'agit d'une spécification autorégressive spatiale (SAR) qui s'écrit sous forme matricielle :

$$Y = \rho WY + X\beta + u \quad (3-8)$$

Où Y est le vecteur des dépenses publiques, X une matrice des caractéristiques observables des communes, β le vecteur des paramètres associés à X et u représente le vecteur des termes d'erreurs avec les propriétés usuelles.

W est la matrice de pondération spatiale, de dimensions $N \times N$ traduisant les interactions entre les communes. Les éléments w_{ij} permettent d'introduire l'hypothèse théorique d'interaction stratégique horizontale : le niveau de dépenses d'équipements d'une commune i est influencé par celles des autres communes j voisines. En reprenant les caractéristiques usuelles des matrices de pondération spatiale, le degré d'influence entre deux communes est supposé décroissant avec la distance qui les sépare. L'intensité des interactions sera d'autant plus forte que la distance qui sépare les communes est faible. Cette hypothèse traduit de manière appropriée le fait que les dépenses d'équipement vont permettre d'améliorer le cadre de vie des habitants d'une commune, ce qui sera d'autant plus visible par les habitants des communes voisines les plus proches. Pour faciliter l'interprétation des paramètres d'interdépendances spatiaux issus de modèles différents, la matrice W est généralement standardisée en ligne.

Chaque élément de la ligne de la matrice est divisé par la somme totale des éléments de la ligne. L'élément de la matrice standardisée, ou bien le poids spatial, est donné par :

$$\tilde{w}_{ij} = \frac{w_{ij}}{\sum_j w_{ij}}$$

ρ est donc le paramètre spatial qui va indiquer la présence ou non des externalités horizontales. Sous l'hypothèse de standardisation de W , ρ mesure l'influence moyenne des choix des communes voisines j sur le choix de la commune i . S'il est non significatif ou nul, cela implique une absence d'interdépendances des dépenses publiques entre les communes.

En revanche, si le paramètre ρ est significatif et positif, cela signifie l'existence d'interdépendances des dépenses publiques entre les communes. Cela implique que la commune i adopte les mêmes comportements en matière de dépenses publiques que ses communes voisines. Par exemple, si ses communes voisines augmentent leurs niveaux de dépenses alors la commune i , procède également à la hausse de ses dépenses. Au contraire, si le paramètre ρ est significatif et négatif, alors l'interdépendance entre les communes est opposée : la commune i adopte un comportement opposé à ceux de ses voisines en matière de dépenses publiques. Si les communes voisines augmentent leurs niveaux de dépenses, alors la commune i procède à la réduction de ses dépenses.

Si la matrice $(I - \rho W)$ est non singulière, l'équation (3-8) peut être réécrite de la manière suivante :

$$\begin{aligned} Y &= (I - \rho W)^{-1} X \beta + (I - \rho W)^{-1} u \\ &= (I - \rho W)^{-1} X \beta + u^* ; \text{ avec } u^* = (I - \rho W)^{-1} u \end{aligned} \quad (3-9)$$

Cette spécification montre la corrélation entre la variable Y et le terme d'erreur u ce qui implique que les estimateurs MCO sont biaisés et non convergents. Le modèle SAR peut être estimé par le maximum de vraisemblance.

Par conséquent, lorsque les interdépendances existent et sont ignorées, l'estimation de l'effet causal est biaisée. *Toutes choses égales par ailleurs*, un effet de traitement biaisé conduit à des conclusions erronées sur les effets réels de la politique et, par conséquent, des décisions politiques fondées sur des résultats biaisés. L'évaluation *ex ante* des comportements des communes impliquées dans le projet Grand Paris doit alors tenir compte de ces effets.

3.3 Modélisation *Difference-in-Difference* du projet Grand Paris

3.3.1. Effet d'anticipation et interdépendances entre les communes

Le projet Grand Paris a été lancé en juin 2007 lors du discours du Président de la République Française sur l'avenir de la région capitale. Ce projet se matérialise sous sa forme actuelle en 2010 par la création de la Société du Grand Paris par la loi du 3 juin 2010. La période 2007 - 2010 a été marquée par des consultations, réalisation d'études, la création d'un secrétariat d'Etat pour le développement de la région (*cf.* Annexe A) jusqu'à l'officialisation des territoires stratégiques et des Contrats de Développement Territorial.

C'est pourquoi, nous supposons que les élus locaux, compte tenu des enjeux, des retombées du projet et de leurs niveaux d'information sur le projet, étaient susceptibles de développer des comportements d'anticipation sur plusieurs aspects notamment le niveau des dépenses d'investissement des communes. On s'attend aussi à ce que l'avènement du projet Grand Paris affecte différemment les communes concernées par rapport aux communes non concernées. Ainsi, l'effet d'anticipation induit par le projet Grand Paris est approximé par l'effet des CDT sur les dépenses d'équipement des communes.

La disponibilité des données pour l'ensemble des communes à différentes dates, nous conduit à l'estimation de l'effet du projet par la méthode *Difference-in-Difference*, notée DID que nous étendons à un cadre spatialisé pour tenir compte des interdépendances entre les communes. Le modèle développé est un modèle *Difference-in-Difference* spatialisé, noté SDID.

3.3.2. Modélisation *Difference-in-Difference* standard

Nous appellerons traitement la loi n° 2010-597 du 3 juin 2010 relative au Grand Paris et plus précisément son article 21 relatif aux CDT. Nous nous intéressons aux décisions des communes franciliennes en termes de dépenses d'équipement, en contrôlant leur appartenance ou non à un CDT.

L'estimateur de DID consiste à mesurer la différence des résultats observés en contrôlant la différence de statut des individus, traité ou non traité, et la différence de période concernée, après et avant le traitement. L'avantage de la méthode DID est qu'elle permet de supprimer les biais liés aux effets fixes des deux groupes, mais aussi les biais liés aux effets temporels dans le groupe de traitement (Imbens et Wooldridge, 2009, p.67).

La méthode consiste, d'abord, à faire une première différence des résultats observés après et avant traitement pour chaque groupe et, ensuite, à faire la différence des différences, des deux groupes. La première différence élimine les effets fixes de chaque groupe et la seconde différence élimine les effets temporels communs aux deux groupes.

Le modèle *Difference-in-Difference* peut être exprimé sous la forme suivante :

$$y_i = \alpha + x_i'\beta + \theta D_i + \gamma \Gamma_i + \delta(D_i * \Gamma_i) + u_i \quad (3-10)$$

y_i est le résultat observé pour la commune (les dépenses d'équipement), x_i un vecteur des caractéristiques observables de la commune (population, potentiel fiscal,...). D_i est une variable indicatrice qui est égale à 1 si la commune est dans un CDT et 0 si la commune est hors CDT.

Nous observons les données sur deux dates en T et T' . Γ_i est une variable indicatrice qui est égale à 1 si la date est postérieure au traitement et 0 si elle correspond à une situation avant le traitement. Dans la mesure où le traitement ici n'est pas encore réalisé au sens où la politique du Grand Paris n'est pas achevée, les marqueurs temporels définissant l'avant et l'après vont correspondre à des événements clés de l'avancement du projet (cf. 3.4).

α est la constante et β est le vecteur des paramètres associés au vecteur x_i .

θ représente l'effet de groupe, γ représente l'effet temporel qui est commun aux deux groupes, **le paramètre δ mesure le vrai effet du traitement.**

- Pour les communes CDT

Après le traitement, c'est-à-dire $\{D = 1 \text{ et } \Gamma = 1\}$, l'équation des résultats observés en termes de dépenses d'équipements notée Y_1^T est donnée par :

$$Y_1^T = \alpha + X_1\beta + \theta + \gamma + \delta + u_{1T}$$

Avec une espérance $\bar{Y}_1^T = E[Y_1^T] = \alpha + X_1\beta + \theta + \gamma + \delta$

Avant le traitement, $\{D = 1 \text{ et } \Gamma = 0\}$, l'équation des résultats sur les dépenses d'équipement notée Y_0^T s'écrit de la manière suivante :

$$Y_0^T = \alpha + X_0\beta + \theta + u_{0T}$$

Avec une espérance $\bar{Y}_0^T = E[Y_0^T] = \alpha + X_0\beta + \theta$

La première différence des espérances qui permet d'éliminer l'effet de groupe « CDT », est alors donnée par :

$$\begin{aligned}
 \pi_T &= \bar{Y}_1^T - \bar{Y}_0^T \\
 &= E[Y_1^T] - E[Y_0^T] \\
 &= (\alpha + X_1\beta + \theta + \gamma + \delta) - (\alpha + X_0\beta + \theta) \\
 &= (\alpha - \alpha) + (X_1\beta - X_0\beta) + (\theta - \theta) + \gamma + \delta \\
 &= \Delta X_{10}\beta + \gamma + \delta
 \end{aligned} \tag{3-11}$$

- Pour les communes hors CDT

Après le traitement, $\{D = 0 \text{ et } \Gamma = 1\}$, l'équation des résultats sur les dépenses d'équipement notée Y_1^C s'écrit de la manière suivante:

$$Y_1^C = \alpha + X_1\beta + \gamma + u_{1c}$$

Avec une espérance $\bar{Y}_1^C = E[Y_1^C] = \alpha + X_1\beta + \gamma$

Avant le traitement, $\{D = 0 \text{ et } \Gamma = 0\}$, l'équation des résultats sur les dépenses d'équipement notée Y_0^C s'écrit de la manière suivante :

$$Y_0^C = \alpha + X_0\beta + u_{0c}$$

Avec une espérance $\bar{Y}_0^C = E[Y_0^C] = \alpha + X_0\beta$

La première différence des espérances qui permet d'éliminer l'effet de groupe « Hors CDT » est alors donnée par :

$$\begin{aligned}
 \pi_C &= \bar{Y}_1^C - \bar{Y}_0^C \\
 &= E[Y_1^C] - E[Y_0^C] \\
 &= (\alpha + X_1\beta + \gamma) - (\alpha + X_0\beta) \\
 &= (\alpha - \alpha) + (X_1\beta - X_0\beta) + \gamma \\
 &= \Delta X_{10}\beta + \gamma
 \end{aligned} \tag{3-12}$$

- **L'estimateur de l'effet de traitement** est obtenu à travers la seconde différence entre les équations (3-11) et (3-12). Elle permet également d'éliminer également l'effet temporel.

$$\begin{aligned}
 \pi_{DID} &= (\bar{Y}_1^T - \bar{Y}_0^T) - (\bar{Y}_1^C - \bar{Y}_0^C) \\
 &= \pi_T - \pi_C \\
 &= (\Delta X_{10}^T \beta + \gamma + \delta) - (\Delta X_{10}^C \beta + \gamma) \\
 &= (\Delta X_{10}^T \beta - \Delta X_{10}^C \beta) + (\gamma - \gamma) + \delta \\
 &= (\Delta X_{10}^T \beta - \Delta X_{10}^C \beta) + \delta
 \end{aligned} \tag{3-13}$$

Généralement, l'effet de traitement δ (équation 3-12) peut être estimé en utilisant les méthodes des moindres carrés sur la fonction de régression de la différence de résultats observés :

$$\Delta Y_{TT'} = \Delta X_{TT'} \beta + \delta \Delta D_{TT'} + \Delta u_{TT'} \tag{3-14}$$

- Où T et T' représentent respectivement les périodes après et avant le traitement.
- $\Delta Y_{TT'}$ est le vecteur de la variable dépendante de dimension $(N \times 1)$, égal à la différence des résultats observés après et avant le traitement, c'est-à-dire l'évolution des dépenses d'équipement des communes.
- $\Delta X_{TT'}$ est une matrice des variables explicatives de dimension $(N \times k)$, égale à la différence des variables explicatives observées après et avant le traitement, c'est-à-dire l'évolution des caractéristiques des communes.
- $\Delta u_{TT'}$ est le nouveau vecteur des termes d'erreur de dimension $(N \times 1)$.
- $\Delta D_{TT'}$ est le vecteur de la variable de traitement représentant l'évolution des statuts des communes. Avant le traitement (donc avant 2010) l'ensemble des communes sont considérées non traités et après le traitement seulement les communes CDT ont changé de statut ce qui fait que la variable est binaire ($\Delta D_{TT'} = \{0,1\}$).
- Les paramètres β et δ sont définis comme précédemment.

Cependant, l'estimation de l'effet de traitement à partir de cette formulation équation (3-14) ne prend pas en compte les effets interactions entre les communes. Pour intégrer les interdépendances entre les communes, nous proposons l'estimation de l'effet de traitement correspondant, avec le modèle *Difference-in-Difference spatial*.

3.3.3. Le modèle *Difference-in-Difference* spatial (SDID)

Nous supposons que le niveau de dépenses d'équipement d'une commune dépend de ses caractéristiques mais aussi du niveau de dépenses d'équipement des communes voisines.

Le modèle *Difference-in-Difference* prenant en compte le niveau de dépenses des communes voisines peut être exprimé sous la forme d'un modèle spatial autorégressif SAR. Cette spécification spatiale a été employée par Dubé et *al.* (2014) dans l'estimation, en simple différence, des valeurs immobilières d'appartements ayant fait l'objet d'une transaction ou non. Notre modèle s'étend au cas des doubles différences. L'équation (3–10) devient :

$$y_i = \alpha + \rho \sum_{j \neq i} w_{ij} y_j + x_i' \beta + D_i \theta + \gamma \Gamma_i + \delta_s (D_i * \Gamma_i) + u_i \quad (3-15)$$

Les variables $y_i, x_i, D_i, \Gamma_i, u_i$ et les paramètres $(\alpha, \beta, \theta, \gamma)$ sont définis comme précédemment.

Le nouveau paramètre δ_s représente le vrai effet du traitement tenant compte des effets d'interaction : ρ représente le paramètre mesurant le degré d'interdépendance des dépenses d'équipements entre les communes.

Sous forme matricielle l'équation (3–15) se décline comme suit :

$$Y = \alpha + \rho WY + X\beta + D\theta + \gamma\Gamma + \delta_s(D * \Gamma) + u$$

Pour les communes CDT *après le traitement*, c'est-à-dire $\{D = 1 \text{ et } \Gamma = 1\}$, l'équation spatiale des dépenses d'équipement, notée Y_1^{TS} est donnée par :

$$Y_1^{TS} = \rho WY_1^{TS} + \alpha + X_1^T \beta + \theta + \gamma + \delta_s + u_1$$

Avec une espérance $\bar{Y}_1^{TS} = E[Y_1^{TS}] = \rho WY_1^{TS} + \alpha + X_1^T \beta + \theta + \gamma + \delta_s$

Pour les communes CDT *avant le traitement*, $\{D = 0 \text{ et } \Gamma = 0\}$, l'équation spatiale des dépenses d'équipement, notée Y_0^{TS} est donnée par :

$$Y_0^{TS} = \rho W Y_0^{TS} + \alpha + X_0^T \beta + \theta + u_1$$

Avec une espérance $\bar{Y}_0^{TS} = E[Y_0^{TS}] = \rho W Y_0^{TS} + \alpha + X_0^T \beta + \theta$

Pour le groupe des communes CDT, la première différence des espérances est donnée par :

$$\begin{aligned} \pi_{ST} &= \bar{Y}_1^{TS} - \bar{Y}_0^{TS} \\ &= E[Y_1^{TS}] - E[Y_0^{TS}] \\ &= (\rho W Y_1^{TS} + \alpha + X_1^T \beta + \theta + \gamma + \delta_S) - (\rho W Y_0^{TS} + \alpha + X_0^T \beta + \theta) \\ &= (\alpha - \alpha) + (\rho W Y_1^{TS} - \rho W Y_0^{TS}) + (X_1^T \beta - X_0^T \beta) + (\theta - \theta) + \gamma + \delta_S \\ &= (\rho W \Delta Y_{10}^{TS}) + (\Delta X_{10}^T \beta) + \gamma + \delta_S \end{aligned} \quad (3-16)$$

Pour les communes hors CDT **après le traitement**, $\{D = 0 \text{ et } \Gamma = 1\}$, l'équation spatiale des dépenses d'équipement, notée Y_1^{CS} est donnée par :

$$Y_1^{CS} = \rho W Y_1^{CS} + X_1^C \beta + \alpha + \gamma T + u_1$$

Avec une espérance $\bar{Y}_1^{CS} = E[Y_1^{CS}] = \rho W Y_1^{CS} + X_1^C \beta + \alpha + \gamma$

Pour les communes hors CDT **avant le traitement**, $\{D = 0 \text{ et } \Gamma = 0\}$, l'équation spatiale des dépenses d'équipement, notée Y_0^{CS} est donnée par :

$$Y_0^{CS} = \rho W Y_0^{CS} + \alpha + X_0^C \beta + u_0$$

Avec une espérance $\bar{Y}_0^{CS} = E[Y_0^{CS}] = \rho W Y_0^{CS} + X_0^C \beta + \alpha$

Pour le groupe des communes hors CDT, la première différence des espérances est donnée par :

$$\begin{aligned} \pi_{SC} &= \bar{Y}_1^{CS} - \bar{Y}_0^{CS} \\ &= E[Y_1^{CS}] - E[Y_0^{CS}] \\ &= (\rho W Y_1^{CS} + \alpha + X_1^C \beta + \gamma) - (\rho W Y_0^{CS} + X_0^C \beta + \alpha) \\ &= (\alpha - \alpha) + (\rho W E Y_1^{CS} - \rho W Y_0^{CS}) + (X_1^C \beta - X_0^C \beta) + \gamma \\ &= (\rho W \Delta Y_{10}^{CS}) + (X_{10}^C \beta) + \gamma \end{aligned} \quad (3-17)$$

Enfin, à partir des équations (3-16) et (3-17), l'estimateur de l'effet de traitement par DID-Spatial (SDID) est donné par :

$$\begin{aligned}
 \pi_{SDID} &= \pi_{ST} - \pi_{SC} \\
 &= (\rho W \Delta Y_{10}^{TS} + \Delta X_{10}^{TS} \beta + \gamma + \delta_S) - (\rho W \Delta Y_{10}^{CS} + \Delta X_{10}^C \beta + \gamma) \\
 &= (\rho W \Delta Y_{10}^{TS} - \rho W \Delta Y_{10}^{CS}) + (\Delta X_{10}^T \beta - \Delta X_{10}^C \beta) + (\gamma - \gamma) + \delta_S \\
 &= \rho W (\Delta Y_{10}^{TS} - \Delta Y_{10}^{CS}) + (\Delta X_{10}^T \beta - \Delta X_{10}^C \beta) + \delta_S
 \end{aligned} \tag{3-18}$$

L'estimateur de l'effet de traitement avec la prise en compte des interdépendances spatiales δ_S , peut être estimé à partir d'un modèle autorégressif spatial SAR suivant :

$$\Delta Y_{TT'} = \rho W \Delta Y_{TT'} + \Delta X_{TT'} \beta + \delta_S \Delta D_{TT'} + \Delta u_{TT'} \tag{3-19}$$

Si la matrice $(I - \rho W)$ est non singulière, l'équation (3-19) peut être réécrite sous la forme réduite suivante :

$$\Delta Y_{TT'} = (I - \rho W)^{-1} [\Delta X_{TT'} \beta + \delta_S \Delta D_{TT'} + \Delta u_{TT'}] \tag{3-20}$$

Le Tableau 3-1 récapitule les différents estimateurs et permet de montrer que sous certaines conditions l'estimateur SDID de l'effet de traitement est égal à l'effet de traitement avec l'estimateur DID. En effet, si le paramètre ρ est non significatif, donc en l'absence d'interdépendances, alors $\pi_{SDID} = \pi_{DID}$. En revanche si le paramètre ρ est significatif alors $\pi_{SDID} \neq \pi_{DID}$, par conséquent l'estimateur π_{DID} par les moindres carrés ordinaires est biaisé et ne traduit plus correctement l'effet réel de la politique.

Tableau 3-1 : Tableau récapitulatif des estimateurs en *DID* et *DID-spatial*

DID	Prétraitement	Post traitement	Différence
Communes CDT	$\alpha + X_0^T \beta + \theta$	$\alpha + X_1^T \beta + \theta + \gamma + \delta$	$\Delta X_{10}^T \beta + \gamma + \delta$
Communes hors CDT	$\alpha + X_0^C \beta$	$\alpha + X_1^C \beta + \gamma$	$\Delta X_{10}^C \beta + \gamma$
Sans interdépendances			$(\Delta X_{10}^T - \Delta X_{10}^C) \beta + \delta$

DID-spatial	Prétraitement	Post traitement	Différence
Communes CDT	$\rho W Y_0^{TS} + \alpha + X_0^T \beta + \theta$	$\rho W Y_1^{TS} + \alpha + X_1^T \beta + \theta + \gamma + \delta_s$	$\rho W \Delta Y_{10}^{TS} + \Delta X_{10}^T \beta + \gamma + \delta_s$
Communes hors CDT	$\rho W Y_0^{CS} + X_0^C \beta + \alpha$	$\rho W Y_1^{CS} + X_1^C \beta + \alpha + \gamma$	$\rho W \Delta Y_{10}^{CS} + \Delta X_{10}^C \beta + \gamma$
Avec interdépendances			$\rho W \Delta (\Delta Y_{10}^{TS} - \Delta Y_{10}^{CS}) + (\Delta X_{10}^T \beta - \Delta X_{10}^C \beta) + \delta_s$

3.4 Le modèle empirique : données et spécification des interactions

L'objectif est d'estimer l'effet d'anticipation des communes sur les dépenses d'équipement par habitant (modèle de concurrence fiscale horizontal) induit par projet du Grand Paris.

3.4.1. Structure de l'échantillon et variables

La variable d'intérêt qui nous intéresse est le niveau de dépenses d'équipement par habitant de la commune. Les dépenses d'équipement sont définies comme étant l'ensemble des dépenses brutes de la commune comprenant « les acquisitions de biens meubles et immeubles, les travaux en cours, les immobilisations incorporelles, les travaux d'investissement en régie et les opérations pour compte de tiers »³⁶. Le choix de cette variable est motivé par le fait que les dépenses d'équipement constituent une variable stratégique (Wildasin, 1988) dans la compétition fiscale entre les communes, et par conséquent, elles permettent d'introduire la possibilité d'externalités horizontales dans l'estimation du traitement. D'autre part, elles représentent une part importante des dépenses d'investissement dans les communes (57,8 % soit près 30 milliards € en 2012³⁷).

Enfin, elles sont en droite ligne des dépenses qui peuvent être impactées par le projet du Grand Paris : le développement de la commune qui peut résulter du projet induira des dépenses nouvelles. L'impact du projet peut également être affecté par le niveau et la qualité *ex ante* des équipements des communes qui cherchent alors à être prête de manière anticipée.

L'échantillon utilisé est le même que celui employé dans l'étude précédente (Chapitre 2). Il comprend 156 communes qui ont adhéré à un des 21 contrats de développement territorial. C'est le groupe des communes CDT. Les 327 autres communes constituent le groupe de contrôle hors CDT. Elles ont été choisies afin de

³⁶<http://www.legifrance.gouv.fr/affichCodeArticle.do?cidTexte=LEGITEXT000006070633&idArticle=LEGIARTI000029968891>, consulté le 27 Septembre 2015

³⁷ Rapport de l'Observatoire des Finances Locales : *Les finances des collectivités locales en 2013*, p.25

respecter une continuité géographique et institutionnelle. L'échantillon des 483 communes couvre ainsi un territoire sans enclave. Ce territoire s'étend au-delà de la limite périphérique des CDT afin de ne pas ignorer l'effet frontière en considérant que le périmètre du Grand Paris se terminerait brusquement aux frontières des CDT périphériques. La continuité institutionnelle est alors captée par ce territoire et les communes connaissent la situation des territoires stratégiques et le projet du GPE. Elles peuvent anticiper les effets des opérations d'aménagement et les comportements stratégiques des autres communes face à ces projets et leurs impacts.

Pour mesurer cet effet, nous disposons d'une base de données fiscales de la DGFIP³⁸ couvrant la période 2001- 2012 et pour l'ensemble des communes de l'Île-de-France. Elle regroupe des informations relatives au budget des communes sur les différentes dépenses et recettes. Les caractéristiques socioéconomiques et démographiques des communes sont quant à elles issues des bases de données fournies par l'INSEE. Plus précisément, nous avons utilisé les données « Evolution et structure de la population ».

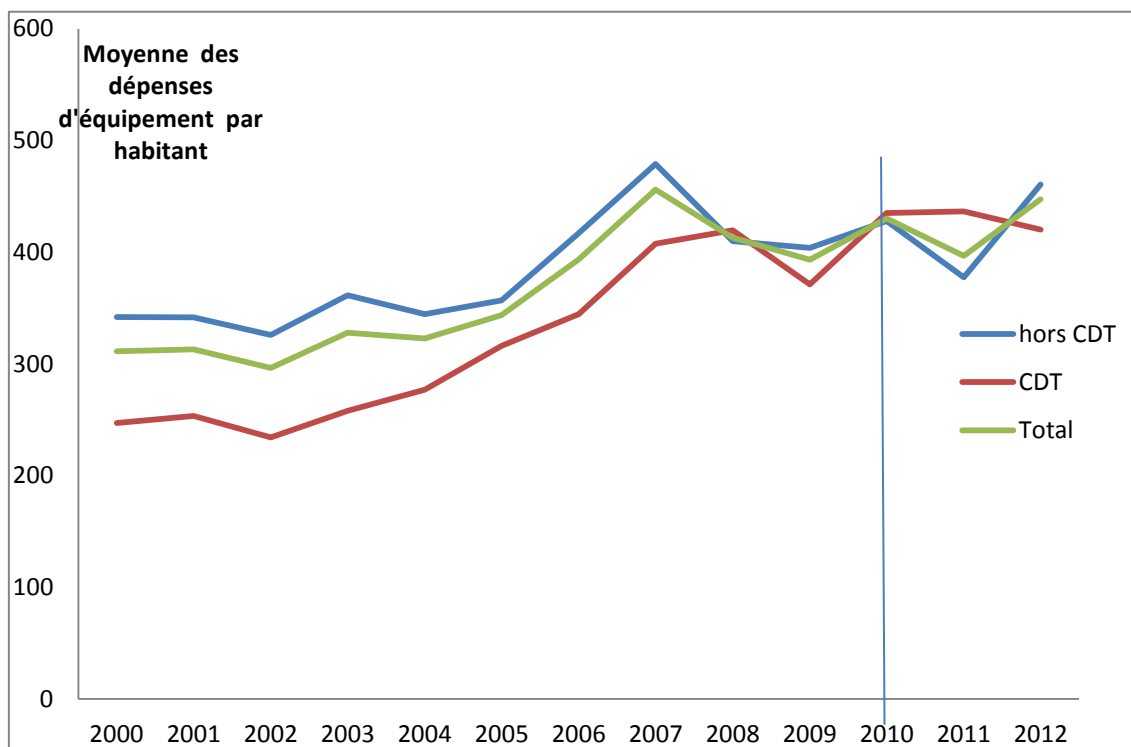
L'analyse de l'évolution des dépenses d'équipement par habitant³⁹ (Figure 3-1 et Tableau 3-2) montre des situations différentes dans le temps et selon le statut des communes.

Pour l'ensemble des communes, sur la période 2001-2012 (Figure 3-1), on observe deux périodes : une tendance à la croissance des dépenses d'équipement de 2001 à 2007 puis ensuite une période plus perturbée jusqu'en 2012. On observe également une baisse des dépenses d'équipement par habitant en 2002 et en 2008. Ces baisses coïncident avec les années post-électorales qui sont souvent marquées par un repli des investissements en général. En plus de ce phénomène post électoral, la baisse de 2002 pourrait s'expliquer par la création, la même année, des structures intercommunales qui favorisent la réalisation des investissements à l'échelon intercommunal, alors que la baisse plus forte observée en 2008 serait au-delà de l'après 2007, la conséquence de la crise économique et financière qui touche le pays.

³⁸ Direction générale des Finances publiques (DGFIP)

³⁹ Les dépenses d'équipement par habitant sont exprimées en € constant.

Figure 3-1 : Evolution des dépenses d'équipement des communes



Source : Direction générale des Finances publiques 2000-2012 (en € constant)

Compte tenu de ces évolutions et des événements marquants du projet du Grand Paris, nous choisissons d'analyser les effets d'anticipation des communes sur trois périodes.

- La première période 2001-2008 correspond à l'effet d'anticipation sur l'annonce de la « Vision pour la Région Capitale » du 26 juin 2007. Elle coïncide avec la phase de croissance tendancielle des dépenses d'équipement.

- La deuxième période 2008-2012 correspond à l'effet d'anticipation sur les dispositions officielles relatives aux CDT de l'article 21 de la loi du 3 juin 2010 sur le Grand Paris. Elle coïncide avec la phase de reprise de croissance des dépenses d'équipement. Le choix de 2012 permet d'avoir un décalage budgétaire d'au moins un an pour apprécier l'effet de l'engagement dans les CDT sur les dépenses d'équipement des communes.

- La troisième période 2001-2012 couvre l'ensemble des données disponibles.

Les différentes évolutions sur ces périodes sont données dans le Tableau 3-2. On note, sur l'ensemble de la période 2001-2012 (cf. Tableau 3-2), que les dépenses

d'équipement par habitant des communes CDT sont plus élevées que celles des communes hors CDT, avec une évolution de 167 € en moyenne pour les communes CDT contre 119 € en moyenne pour les communes hors CDT. Les communes CDT dépensent 40% de plus que les communes hors CDT. Pour l'ensemble des communes, les dépenses d'équipement par habitant étaient de 313 € en moyenne, en 2001, contre 447,5 €, en moyenne, en 2012, soit une évolution moyenne de 137 € entre 2001 et 2012. Entre 2001 et 2012, la croissance des dépenses d'équipement a été de 66 % pour les communes CDT contre 35 % pour les communes hors CDT et 43 % sur l'ensemble des communes. Si on s'attache à l'évolution sur les deux sous-périodes 2001-2008 d'une part et 2008-2012, d'autre part, on remarque que les communes CDT ont une évolution plus forte que les communes hors CDT sur la première sous-période (un taux de croissance de leurs dépenses d'équipement de 66 % contre 20 % pour les communes hors CDT) alors que sur la seconde sous-période, les communes CDT enregistrent une stagnation de leurs dépenses et que les communes hors CDT maintiennent une certaine croissance (12%).

Tableau 3-2 : Statistiques descriptives des dépenses d'équipement

Communes	CDT	Hors CDT	Ensemble
Résultat	Moyenne (écart-type) <i>Tx de croissance</i> (%)	Moyenne (écart-type) <i>Tx de croissance</i> (%)	Moyenne (écart-type) <i>Tx de croissance</i> (%)
Dépenses d'équipement par habitant (en €) en 2001	253,57 (10,974)	341,67 (32,97)	313,21 (22,67)
Dépenses d'équipement par habitant (en €) en 2008	419,54 (26,8)	409,92 (30,35)	413,03 (22,28)
Dépenses d'équipement par habitant (en €) en 2012	420,21 (22,263)	460,58 (45,33)	447,54 (31,51)
Evolution des dépenses d'équipement par habitant 2001-2012	166,64 (21,04) 66%	118,91 (37,86) 35%	134,33 (26,52) 43%
Evolution des dépenses d'équipement par habitant 2001-2008	165,97 (26,04) 65%	68,25 (26,437) 20%	99,82 (19,869) 32%
Evolution des dépenses d'équipement par habitant 2008-2012	0,67 (26,4) 0%	50,7 (34,5) 12%	34,5 (-24,8) 8%

Les principaux déterminants des dépenses d'équipement par habitant que nous retenons sont issus de la littérature consacrée à ce sujet (Case et al., 1993; Solé-Ollé, 2006) et sont par ailleurs issus de la procédure *stepwise* de sélection des variables. Il s'agit des variables relatives à la richesse des communes, mesurée à l'aide du potentiel fiscal (la base de la taxe à l'habitation, la base de la taxe sur le foncier bâti et la base de la taxe sur le foncier non bâti⁴⁰) et des variables démographiques. Nous présentons quelques statistiques descriptives de ces variables reportées dans les Tableau 3-3a et Tableau 3-3b.

Entre 2001 et 2012, la base de la taxe à l'habitation par habitant est passée en moyenne de 1272 € à 1673 €, soit une hausse de 32 % correspondant à une

⁴⁰ Ces différentes bases correspondent à la valeur locative cadastrale des propriétés correspondantes. Le montant des recettes fiscales issues des 3 taxes locales est obtenu en multipliant les bases par les taux d'imposition fixés, par les communes.

augmentation de 400 € par habitant en moyenne. Cette augmentation est identique dans les communes CDT ou hors CDT (32%) mais pour des valeurs initiales plus fortes dans les communes CDT. Sur les deux sous-périodes, les évolutions sont similaires en % selon le type de communes. Pour la base de la taxe sur le foncier bâti par habitant, la hausse moyenne est de 583 €, soit 39 % d'augmentation en moyenne. Là encore, la croissance est du même ordre pour les communes CDT ou hors CDT, mais le rythme diffère sur les deux sous-périodes : On enregistre une croissance plus faible pour les communes CDT entre 2001 et 2008 (25 % contre 30 % pour les communes hors CDT) alors que pour les communes hors CDT la croissance est plus faible entre 2008 et 2012 (7 % contre 11 % pour les communes CDT). Quant à la base de la taxe sur le foncier non bâti par habitant, elle est passée de 17,36 €, en moyenne, en 2001, à 15,19 € par habitant en moyenne, en 2012, soit une baisse de 12,5 %. Cette baisse est forte en % pour les communes hors CDT (-13 % contre -7 % pour les communes CDT). La baisse est observée sur les deux sous-périodes et elle est forte dans les communes hors CDT sur la période 2001-2008 (12 % contre 2 % pour le communes CDT).

Tableau 3-3a : Statistiques descriptives des caractéristiques des communes
Données fiscales

Communes	CDT	Hors CDT	Ensemble
Variables	Moyenne	Moyenne	Moyenne
	(écart-type) <i>Tx de croissance</i> (%)	(écart-type) <i>Tx de croissance</i> (%)	(écart-type) <i>Tx de croissance</i> (%)
Base de la taxe d'habitation par habitant (TH/en €) en 2001	1213,756 (30,189)	1299,481 (24,535)	1271,794 (19,330)
Base de la taxe d'habitation par habitant (TH/en €) en 2008	1475,103 (35,285)	1606,113 (29,549)	1563,799 (23,171)
Base de la taxe d'habitation par habitant (TH/en €), en 2012	1587,152 (37,986)	1713,862 (30,156)	1672,937 (23,949)
Evolution de la TH 2001-2012	373,396 (12,764) 31%	414,381 (8,919) 32%	401,143 (7,356) 32%
Evolution de la TB 2001-2008	261,35 (8,393) 22%	306,63 (7,987) 24%	292,01 (6,121) 23%
Evolution de la TFB 2008-2012	203,52 (39,600) 8%	130,12 (50,170) 8%	153,820 (36,300) 8%
Base de la taxe sur le foncier bâti par habitant (TFB/en €), en 2001	1442,895 (96,595)	1535,910 (185,350)	1505,870 (129,240)
Base de la taxe sur le foncier bâti par habitant (TFB/en €), en 2008	1799,229 (122,275)	1999,44 (285,320)	1934,780 (197,100)
Base de la taxe sur le foncier bâti par habitant (TFB/en €), en 2012	2002,744 (150,269)	2129,560 (309,380)	2088,600 (214,900)
Evolution de la TFB 2001-2012	559,850 (70,670) 39%	593,700 (170,700) 39%	582,730 (117,720) 39%
Evolution de la TFB 2001-2008	356,334 (36,52) 25%	463,53 (129,567) 30%	428,91 (88,492) 28%
Evolution de la TFB 2008-2012	203,52 (39,600) 11%	130,12 (50,170) 7%	153,820 (36,300) 8%
Base de la taxe sur le foncier non bâti par habitant (TFNB/en €) en 2001	7,790 (1,180)	21,930 (2,350)	17,360 (1,660)
Base de la taxe sur le foncier non bâti par habitant (TFNB/en €) en 2008	7,663 (0,998)	19,4 (1,980)	15,610 (1,400)
Base de la taxe sur le foncier non bâti par habitant (TFNB/en €), en 2012	7,242 (0,796)	18,970 (1,820)	15,190 (1,280)
Evolution de la TFNB 2001-2012	-0,549 (0,730) -7%	-2,951 (0,734) -13%	-2,175 (0,552) -13%
Evolution de la TFNB 2001-2008	-0,127 (0,750) -2%	-2,53 (0,610) -12%	-1,75 (0,480) -10%
Evolution de la TFNB 2008-2012	-0,4203 (0,5127) -5%	-0,4302 (0,365) -2%	-0,427 (0,297) -3%

Source: base de données DGFIP 2000-2012

Les déterminants relatifs aux caractéristiques démographiques des communes retenus sont : la part de la population âgée de moins de 15 ans, la part de la population âgée de 75 ans et plus, ainsi que la densité de la commune. Ces variables permettent de capter l'impact de la structure de la population qui conditionne en partie les dépenses d'équipements des communes.

Les communes caractérisées par une population relativement jeune vont axer leurs investissements sur l'éducation, le sport et les loisirs, par exemple, alors que dans les communes dont la proportion des personnes âgées est relativement importante, les investissements seront plus orientés vers les services d'aide à la personne, les services d'hébergement, les espaces verts,...

La population totale de l'ensemble des communes est passée de 16 000 à 17 100 habitants en moyenne entre 2001 et 2012 soit un accroissement de 7 %. La répartition de la population montre que les communes CDT sont deux fois plus peuplées en moyenne que les communes hors CDT. En 2001, les communes CDT comptaient en moyenne 26 400 habitants contre 11 000 habitants en moyenne pour les communes hors CDT et en 2012, elle était de 28 400 habitants en moyenne pour les communes CDT contre 11 700 habitants en moyenne pour les communes hors CDT. Sur la période 2001-2012, la population des CDT a augmenté de 8 % et celle des communes hors CDT de 6 %. L'accroissement de la population implique naturellement une augmentation de la demande des services dans les communes.

On observe également une hausse de la densité au km² passant de 13 602 habitants au km², en 2001, à 14 556 habitants au km², en 2012 en moyenne, soit une augmentation de 9 %. Sur la période 2001-2012, la densité a augmenté de presque de 10 % dans les communes CDT contre 8 % dans les communes hors CDT. La croissance de la population s'est donc effectuée avec une densification de l'espace. Cette forte densité peut avoir, d'une part, un effet positif sur les dépenses des communes, car elle peut permettre à la commune de réaliser des économies d'échelle, en réduisant certaines dépenses, par exemple la collecte des ordures ménagères. D'autre part, cette densification peut induire des dépenses supplémentaires du fait de la croissance de la population (scolaire, bibliothèques, sport et loisirs,...), des services d'assistance, maisons de retraites,... Entre 2001 et 2012, la part de la population de moins de 15 ans a faiblement évolué, toutefois cette part est à la baisse : un peu moins fortement dans les

communes CDT que dans les communes hors CDT. En revanche la part de la population âgée a augmenté, aussi bien pour les communes CDT que hors CDT. Elle est passée de 4 %, en moyenne, en 2001, à 6 % en moyenne, en 2012 avec une hausse relativement plus forte pour les communes hors CDT (+1,8%) que pour les communes CDT (+1,5%).

Tableau 3-3b : Statistiques descriptives des caractéristiques des communes
Données démographiques

Communes	CDT	Hors CDT	Ensemble
Variables	Moyenne (écart-type) <i>Tx de croissance (%)</i>	Moyenne (écart-type) <i>Tx de croissance (%)</i>	Moyenne (écart-type) <i>Tx de croissance (%)</i>
Densité (km ²) en 2001	15508,604 (380,186)	12825,348 (229,669)	13691,989 (205,991)
Densité (km ²) en 2008	16593,662 (393,132)	13709,512 (235,903)	14641,040 (212,860)
Densité (km ²) en 2012	16889,639 (395,678)	13886,570 (238,330)	14856,505 (215,317)
Evolution de la densité 2001-2012	9%	8%	9%
Evolution de la densité 2001-2008	7%	7%	7%
Evolution de la densité 2008-2012	2%	1%	1%
Part de la population moins de 15 ans, en 2001	0,219 (0,003)	0,211 (0,002)	0,214 (0,001)
Part de la population moins de 15 ans, en 2008	0,212 (0,002)	0,204 (0,001)	0,207 (0,001)
Part de la population moins de 15 ans, en 2012	0,213 (0,002)	0,203 (0,001)	0,206 (0,001)
Evolution de la population jeune 2001-2012	-0,6%	-0,8%	-0,8%
Evolution de la population jeune 2001-2008	-0,7%	-0,7%	-0,7%
Evolution de la population jeune 2008-2012	+0,1%	-0,1%	-0,1%
Part de la population de 75 ans et plus, en 2001	0,042 (0,001)	0,043 (0,001)	0,043 (0,001)
Part de la population de 75 ans et plus, en 2008	0,053 (0,002)	0,057 (0,001)	0,056 (0,001)
Part de la population de 75 ans et plus, en 2012	0,057 (0,002)	0,061 (0,001)	0,060 (0,001)
Evolution de la population senior 2001-2012	+1,5%	+1,8%	+1,7%
Evolution de la population senior 2001-2008	+1,1%	+1,4%	+1,3%
Evolution de la population senior 2008-2012	+0,4%	+0,4%	+0,4%

Sources : Insee, *Evolution et structures de la population, 2001, 2008, 2012.*

3.4.2. Les structures d'interactions : les matrices de voisinage

Pour prendre en compte les interdépendances des dépenses d'équipement des communes, nous avons considéré deux types de matrices de voisinage W . La matrice géographique est basée sur la distance euclidienne entre deux communes. La matrice « économique » est basée sur les échanges flux domicile-travail entre les communes.

La dimension géographique est importante pour tenir compte des effets de débordement géographique traditionnels. L'équipement d'une commune sera plus facilement accessible aux habitants des autres communes proches que des communes plus éloignées. On considère, pour chaque commune i , un voisinage composé de j communes situées dans un rayon de 10 km autour de la commune i . Les éléments w_{ij} de la matrice W sont définis de la manière suivante et par construction $w_{ii} = 0$.

$$w_{ij} = \begin{cases} 1, & \text{si } \text{dist}_{ij} \leq 10 \text{ km} \\ 0, & \text{sinon} \end{cases}$$

dist_{ij} désigne la distance séparant la commune i et la commune j . En moyenne, chaque commune a un voisinage composé de 38 communes. La composition de ces voisinages selon le type de communes est donnée dans le Tableau 3-4. Les communes CDT sont en moyennes entourées de 21 communes CDT et de 10 communes hors CDT. Chaque commune hors CDT a en moyenne 5 voisines CDT et 37 voisines hors CDT.

La matrice économique est construite à partir des flux domicile-travail⁴¹. Nous utilisons la base de données *Mobilités professionnelles* (déplacements domicile - lieu de travail) de l'INSEE de 2012, à l'échelle de la commune. C'est une base qui indique le nombre d'actifs en emploi effectuant des déplacements entre leur lieu de domicile et leur lieu de travail. Les flux ont été obtenus à partir de l'agrégation des données individuelles.

⁴¹ Les bases sur les flux de mobilité des « déplacements domicile-travail » INSEE 2012

Ainsi, nous supposons que deux communes sont voisines si le flux $F_{i \rightarrow j}$ entre elles est non nul. Les éléments w_{ij} de la matrice W sont définis de la manière suivante et par construction $w_{ii} = 0$ i.e $F_{i \rightarrow i} = 0$.

$$w_{ij} = \begin{cases} 1, & \text{si } F_{i \rightarrow j} \neq 0 \\ 0, & \text{sinon} \end{cases}$$

Cette matrice s'appuie sur les interactions issues de l'adéquation entre les emplois et les travailleurs. Elle met en exergue les communes « pourvoyeuses de travailleurs » et les communes d'emplois dans une optique de *spatial matching* et de *skill matching* (appariement sur les compétences), avec des priorités en matière d'équipement différentes. Les communes « pourvoyeuses » pourraient plus axer leurs dépenses d'équipement sur les services résidentiels et à la demande sociale que les communes d'emplois. Pour les flux entre communes associés à des échanges croisés, les demandes d'équipement peuvent également refléter la composition sociale et des CSP des communes. Avec la matrice des flux, chaque commune a 130 communes voisines en moyenne. Il convient de noter que les flux domicile-travail entre deux communes sont assez fréquemment nuls et que le choix de ce seuil, a priori faible, est, cependant, révélateur des connexions entre les communes pour le territoire « Grand Paris » concerné par notre échantillon. Ainsi, pour 465 communes de notre échantillon, la médiane des flux est nulle ce qui signifie que ces communes n'entretiennent aucune relation travail-emploi avec 50 % des communes du territoire « Grand Paris » concerné. Si on regarde le troisième quartile, il est encore nul pour 232 communes : celles-ci sont donc sans relation travail-emploi avec au moins 75 % des communes du territoire « Grand Paris » retenu dans notre échantillon.

Si on regarde la composition moyenne des voisinages selon ce critère, une commune CDT est en contact en moyenne avec 84 communes CDT et 86 hors CDT. Du côté des communes hors CDT, elles sont en contact en moyenne avec 55 communes CDT et 56 communes hors CDT.

Tableau 3-4 : la composition des voisinages des communes

	Type de matrice	
	Géographique	Economique
Ensemble	38	130
CDT-CDT	21	84
CDT-HCDT	10	86
HCDT-HCDT	37	56
HCDT-CDT	5	55

3.5 Résultats des estimations

Nous considérons les trois périodes. Nous estimons dans un premier temps l'effet de l'évolution des caractéristiques des communes sur l'accroissement de leurs dépenses d'équipement, sans tenir compte des interactions spatiales.

$$\Delta Y_{TT'} = \Delta X_{TT'}\beta + \delta\Delta D_{TT'} + \Delta u_{TT'} \quad (3-14)$$

L'équation 3-14 est estimée par les MCO. Nous testons la présence des interactions spatiales et leurs formes et estimons la spécification spatiale (équation (3-19) permettant d'analyser l'impact de l'appartenance des communes à un CDT sur l'évolution de leurs dépenses d'équipement.

$$\Delta Y_{TT'} = \rho W\Delta Y_{TT'} + \Delta X_{TT'}\beta + \delta_S\Delta D_{TT'} + \Delta u_{TT'} \quad (3-19)$$

L'ensemble des variables sont prises en logarithmes. Les équations spatiales sont estimées avec les deux matrices de voisinages pour les périodes 2001-2012 et 2008-2012 et seulement pour la matrice de voisinage géographique pour la période 2001-2008, car nous n'avons pas les données de flux domicile-travail pour l'année 2008.

Les résultats de l'équation (3-14) (colonne [1] à [3] du Tableau 3-5) concernent la période totale 2001-2012 et ses deux sous-périodes 2001-2008 d'une part et 2008-2012, d'autre part. L'effet d'appartenir à un CDT est significatif et positif pour la période totale et égal à 0,017. Le projet Grand Paris semble donc plus impacter le comportement des communes CDT quant à l'évolution de leurs dépenses d'équipement que celui des communes hors CDT. Cependant, si on s'intéresse à un effet *ex ante* de

l'annonce de la « Vision pour une Région Capitale » ou de la constitution des CDT, l'effet n'est plus significatif. De plus, l'effet de traitement estimé pourrait être biaisé si les effets d'interdépendances sont présents.

Tableau 3-5 : L'effet de traitement sans interactions entre les communes

	2001-2012 [1]	2001-2008 [2]	2008-2012 [3]
Variable dépendante : Evolution des dépenses d'équipement (en log)			
Constante	10,240*** (0,206)	-0,230 (0,323)	9,848*** (0,457)
Variables de contrôle			
Log (variation de la densité)	0,010. (0,006)	0,015 (0,014)	0,031 (0,029)
Log (variation de la part des moins de 15 ans)	-0,368* (0,171)	-1,113** (0,426)	-0,345 (0,989)
Log (Variation de la part des 75 ans et plus)	0,961** (0,303)	-1,493. (0,815)	4,34** (1,667)
Log (Variation de la base de la TH/ habitant)	0,000. (0,000)	0,257*** (0,025)	0,031 (0,035)
Log (Variation de la base de la TFB/ habitant)	-0,085*** (0,018)	-0,090*** (0,023)	-0,305*** (0,037)
Log (Variation de la base de la TFNB/ habitant)	-0,147*** (0,036)	1,506*** (0,040)	0,078. (0,042)
Effet du traitement			
CDT (δ)	0,017* (0,008)	0,024 (0,018)	0,012 (0,024)
Diagnostic			
AIC	-1047	-249	22
Log likelihood	532,47	133,30	-1,11
Nombre d'observations	483	483	483

Ecart types entre parenthèses

*Codes significativité : *** significative à 0.01 %, **significativité à 1%, * significativité à 5%, significativité à 10%*

Pour tester la présence de ces effets d'interdépendance, nous procédons à un diagnostic d'autocorrélation spatiale. Nous procédons au test de Moran⁴² (I) sur les résidus de l'équation (3–14). Le test de Moran consiste à tester une hypothèse nulle d'absence d'autocorrélation spatiale sur les résidus de la régression contre une

⁴² L'indice de Moran est défini comme suit : $I = \frac{n}{s} \left\{ \frac{\varepsilon' W \varepsilon}{\varepsilon' \varepsilon} \right\}$, avec ε le vecteur des résidus du modèle MCO, W la matrice de poids spatiaux, n la taille de la population et s facteur de standardisation correspondant à la somme des éléments de W .

hypothèse alternative. Lorsque l'indice (I) est supérieur à la moyenne théorique $E(I)$ ⁴³, cela implique une présence d'autocorrélation spatiale positive, si $I < E(I)$, cela équivaut à une autocorrélation spatiale négative. Pour discriminer entre une spécification autorégressive des erreurs (SEM) ou de la variable endogène (SAR), nous effectuons les tests supplémentaires du Multiplicateur de Lagrange (ML) et leurs versions robustes. Nous appliquons la règle de décision préconisée par Anselin et Florax (1995) qui s'appuie sur la significativité des tests : par exemple, le modèle SEM est choisi si le test RLMERR est significatif alors que le test RLMLAG ne l'est pas ou si lorsque les deux tests sont significatifs, la statistique du test RLMERR est supérieure à la statistique du test RLMLAG. Les résultats montrent que la statistique de Moran est significative avec les deux matrices, pour toutes les spécifications (Tableau 3-6), ce qui implique le rejet de l'hypothèse d'absence d'autocorrélation spatiale des résidus. Les tests robustes RLMERR et RLMLAG suggèrent une correction plus appropriée par l'estimation d'un modèle SAR plutôt que par l'estimation d'un modèle SEM pour les périodes [2001-2012] et [2008-2012]. Le modèle autorégressif ainsi retenu corrobore notre hypothèse d'interdépendances horizontale des dépenses d'équipement entre les communes. En revanche pour la période la plus ancienne « pré annonce » des CDT [2001-2008], aucun test LM n'est significatif. Nous choisissons d'estimer la spécification SAR pour rester conforme à l'hypothèse théorique d'interactions horizontales et tester celle-ci.

⁴³ La moyenne théorique de l'indice $E(I) = -\frac{1}{n-1}$

Tableau 3-6 : Diagnostic de l'interdépendance spatiale

Période	[2001-2012]		[2001-2012]	
	Matrice des flux		Matrice géographique	
	value	p-value	value	p-value
I Moran	-0,019	0,030	-0,025	0,009
LMERR	4,620	0,031	3,540	0,060
RLMERR	0,798	0,371	1,670	0,197
LMLAG	7,450	0,006	5,550	0,018
RLMLAG	3,620	0,057	3,680	0,055

Période	[2001-2008]	
	Matrice des flux	Matrice géographique
	value	p-value
I Moran	0,0339	0,000
LMERR	0,325	0,568
RLMERR	1,391	0,238
LMLAG	1,354	0,244
RLMLAG	2,42	0,12

Période	[2008-2012]		[2008-2012]	
	Matrice des flux		Matrice géographique	
	value	p-value	value	p-value
I Moran	-1,94	0,000	0,012	0,008
LMERR	6,072	0,013	4,048	0,044
RLMERR	1,142	0,285	1,8	0,18
LMLAG	7,722	0,005	5,382	0,020
RLMLAG	2,792	0,095	3,133	0,077

Les résultats des estimations des modèles SAR pour différentes périodes et les deux matrices de voisinage sont reportés dans le Tableau 3-7.

Ce tableau reporte pour les matrices de voisinage les effets marginaux des variables explicatives, à savoir l'effet total d'une variation marginale sur l'évolution des dépenses d'équipement des communes. Ces effets totaux comprennent l'effet de multiplicateur spatial $(I - \rho W)^{-1}$ sur les variables exogènes tel qu'indiqué dans l'équation (3-20)⁴⁴ ci-dessous :

$$\Delta Y_{TT'} = (I - \rho W)^{-1} [\Delta X_{TT'} \beta + \delta_S \Delta D_{TT'} + \Delta u_{TT'}] \quad (3-21)$$

⁴⁴ L'estimation des effets totaux a été réalisée avec le package spdep du logiciel R

Plus précisément, l'effet total du modèle SAR combine l'effet de multiplicateur spatial qui dépend du paramètre d'autocorrélation spatiale ρ et l'effet des caractéristiques des communes qui dépend des estimations des paramètres β et δ .

Plusieurs résultats se dégagent de ces estimations selon la période et les matrices d'interactions.

Concernant l'existence d'effets d'interaction, tout d'abord, nous remarquons qu'il diffère selon la période choisie.

L'estimateur de l'effet d'interaction $\hat{\rho}$ est significativement négative pour les deux matrices et les périodes [2001-2012] et [2008-2012]. Cela implique que les estimateurs du modèle des moindres carrés sont biaisés et inefficients et par conséquent l'effet du traitement estimé $\hat{\delta}$ est biaisé et les conclusions qui en découlent sont biaisées. La significativité du paramètre ρ peut être interprétée du point de vue économique comme l'existence des externalités fiscales horizontales entre les communes : le niveau des dépenses d'équipement par habitant effectué par les communes affectent le niveau des dépenses d'équipement par habitant des communes voisines et réciproquement. Nous détaillerons les implications de ce résultat plus loin.

En revanche pour la période [2001-2008] correspondant à un effet d'anticipation sur l'annonce du Grand Paris, l'effet d'interaction n'est pas significativement différent de 0. On ne peut pas rejeter l'hypothèse nulle d'absence d'interdépendances entre les communes. Les communes ne semblent pas être influencées par le comportement des communes voisines pour leurs dépenses d'équipement.

Nous analysons l'ensemble des résultats pour cette sous-période et les estimations sont données dans la colonne [3] du Tableau 3-7. Nous remarquons d'abord que l'effet de traitement est positif mais non significatif. Les communes n'ont pas non plus de comportement différent au cours de cette période quant à leur décision en matière de dépenses d'équipement qu'elles soient appelées à faire partie d'un CDT ou non en relation avec le futur projet du Grand Paris puis que l'effet de traitement n'est pas significativement différent de 0. La structure démographique des communes influence négativement la variation de leurs dépenses d'équipement. Sur la période, les communes ont connu une diminution de la part de jeunes et une croissance de la part de la population âgée.

Tableau 3-7 : L'effet de traitement avec interactions entre les communes

Périodes	[2001-2012]		[2001-2008]	[2008-2012]	
Structures d'interactions	Matrice des flux [1]	Matrice géographique [2]	Matrice géographique [3]	Matrice des flux [4]	Matrice géographique [5]
Variable dépendante : Evolution des dépenses d'équipement (en log)					
Constante	13,9***	15,5***	0,640	11,8***	14,8***
Variables de contrôle	Effets totaux		SAR	Effets totaux	
Log (Variation de la densité)	0,007●	0,006●	0,014	0,024●	0,018●
Log (Variation de la part des moins de 15 ans)	-0,255*	-0,218*	-1,102**	-0,210●	-0,135●
Log (Variation de la part des 75 ans et plus)	0,670**	0,582**	-1,449●	3,565**	2,577*
Log (Variation de la base de la TH/ habitant)	0,000*	0,000*	0,266***	0,025●	0,020●
Log (Variation de la base de la TFB/ habitant)	-0,062***	-0,052***	-0,094***	-0,245***	-0,180***
Log (Variation de la base de la TFNB/ habitant)	-0,104***	-0,081***	1,514***	0,062●	0,049*
Effet du traitement	Effet totaux		SAR	Effets totaux	
CDT (δ)	0,013*	0,012*	0,027	0,011●	0,008●
Effets d'interaction	SAR		SAR	SAR	
Rho (ρ) (modèle SAR)	-0,419*	-0,605***	-0,115	-0,242●	-0,635***
Diagnostics					
AIC	-1051,2	-1052,5	-248,25	18,733	14,661
Wald stat	4,081*	6,724**	1,782	1,328	6,924**
Log likelihood	535,614	536,236	134,126	0,634	2,670
Nb d'observations	483	483	483	483	483

Niveau de significativité : '●' : 10 % ; '*' : 5 % ; '**' : 1 % ; '***' : 1°%

Dans ce contexte, une diminution de 1 % de la variation de la part des jeunes, produit une évolution à la baisse de 1,1 % des dépenses d'équipement par habitant au cours de la période antérieure au projet Grand Paris. En revanche, une augmentation de 1 % de l'évolution de la part des personnes âgées produit une baisse de 1,4 % de l'évolution des dépenses d'équipement par habitant dans les communes. Une croissance de 1 % du potentiel fiscal des communes influence leurs dépenses d'équipement par habitant, de manière positive pour la taxe à l'habitation (0,26%) et la taxe sur le foncier non bâti (1,5%) et à la baisse pour la taxe sur le foncier bâti (-0,1%).

Les résultats sont différents si on considère la période globale ou la période post annonce du Grand Paris et antérieure à la constitution des CDT. Les comportements des communes sont, d'une part, dépendants des comportements des communes voisines et de manière opposées ($\hat{\rho} < 0$) : lorsque les communes voisines augmentent leurs dépenses d'équipement par habitant, la commune de référence diminue ses dépenses d'équipement par habitant. Cet effet est plus fort pour la proximité géographique, qui assure une sorte de continuité territoriale entre les communes que pour la proximité économique liée aux flux domicile-travail. Sur la période globale comme pour la période post annonce du projet, les effets d'interaction liés à la proximité géographique sont égaux à -0,605 ou -0,635. Un accroissement de 1 % de l'évolution des dépenses d'équipement par habitant dans les communes voisines entraîne une baisse de plus de 0,6 % de l'évolution des dépenses d'équipement par habitant de la commune de référence. De même, les communes qui vont avoir tendance à diminuer leurs dépenses d'équipement par habitant sont en moyenne entourées de communes qui ont tendance à augmenter les leurs. Ce comportement indique que les communes ne souhaitent pas développer leurs dépenses d'équipement si les communes voisines le font. Cela implique que les communes ne cherchent pas à s'imiter et qu'il n'y pas de course au suréquipement ou au sous-équipement. Cela peut signifier au contraire que les communes peuvent raisonner en termes de complémentarité dans l'offre d'équipement motivé par des effets de rationalisation des dépenses et de taille des équipements⁴⁵. Mais cela peut également signifier que les communes peuvent laisser les autres communes améliorer leurs équipements pour pouvoir en bénéficier. Un comportement

⁴⁵ Rappelons que nous nous situons aux niveaux des communes et non de leurs regroupements (EPCI) et que la recherche d'économie d'échelles dans l'offre d'équipement est certainement moindre au niveau communale, puisque les infrastructures mutualisables relèvent plus de l'EPCI que de la commune.

d'exploitation des effets de débordement n'est donc pas à exclure, mais notre application ne permet pas de préciser cela. Pour les effets d'interaction liés à la proximité économique, un accroissement de 1 % de l'évolution des dépenses d'équipement par habitant dans les communes « pourvoyeuses d'emplois » entraîne une baisse de plus de 0,42 % (période 2001-2012) ou de 0,24 % (période 2008-2012) de l'évolution des dépenses d'équipement par habitant de la commune de référence. Les dépenses d'équipement peuvent, en effet, diminuer parce que les actifs utilisent les équipements des communes dans lesquelles ils travaillent.

Avec la prise en compte des effets d'interaction, les résultats montrent que l'effet de traitement correspondant est significatif et positif avec les deux matrices. Il est, suivant la période, égal à 0,013 ou 0,011 avec la matrice des flux et 0,012 ou 0,008 avec la matrice géographique. Ces résultats impliquent que l'appartenance à un CDT a un effet positif sur les dépenses d'équipement des communes. En d'autres termes, l'évolution des dépenses d'équipement par habitant des communes appartenant à un CDT est de 0,8 % à 1,3 % plus importante que celle des communes hors CDT. Surtout, ces effets, significatifs, sont moins importants que ceux fournis par l'estimation MCO, (Tableau 3-5 : 0,017 pour la période globale 2001-2012 ou non significatif pour la période 2008-2012). La non prise en compte des interactions conduit à surestimer l'effet d'appartenance aux CDT ou à négliger cette appartenance. Les résultats MCO sont biaisés et conduisent à de mauvaises interprétations sur l'impact de la constitution du projet Grand Paris sur les dépenses d'équipement des communes.

Pour les effets totaux des variables explicatives sur les dépenses d'équipement par habitant des communes, les résultats pointent vers différentes interprétations. Tout d'abord, que ce soit sur la période globale ou sur la période pendant laquelle l'information sur le Grand Paris et la construction des CDT a beaucoup circulé, les impacts des variables explicatives vont dans le même sens et cela pour les deux matrices d'interactions. L'effet de la densité est positif et plutôt faible sur la période globale alors qu'il joue plus fortement sur la période Grand Paris : Pour les communes ayant une densité 1 % plus élevée, l'évolution des dépenses d'équipement par habitant a été 0,024 % ou 0,018 % plus forte sur la période post annonce du Grand Paris. La composition de la population joue négativement lorsqu'il s'agit des jeunes et très

fortement et positivement lorsqu'il s'agit de la population âgée. Le vieillissement de la population pèse sur l'évolution des dépenses d'équipement des communes et surtout sur la période récente. La variation de 1 % de la part de la population âgée induirait une variation de 3,5 % ou 2,6 % des dépenses d'équipement par habitant pour les communes concernées. Les effets totaux sont inférieurs aux valeurs estimées par les MCO, ce qui corrobore l'effet de surestimation associé à l'absence de prise en compte des interactions entre les communes.

Concernant les taxes, l'effet à la baisse est plus fort sur la période récente : une amélioration de la base locative cadastrale sur les propriétés bâties contribue à une évolution à la baisse des dépenses d'équipement des communes. La variation à la hausse de 1 % de la base de la taxe sur le foncier bâti entraîne une baisse des dépenses d'équipement par habitant de 0,24 % avec la matrice des flux et 0,18 % avec la matrice géographique.

L'accroissement de la base locative au contraire induit une hausse des dépenses d'équipement qui pourrait être liée à la croissance de la population (plus de logements habités). La variation à la hausse de 1 % de la base de la taxe sur le foncier bâti se traduit par une de presque de 0,02 % des dépenses d'équipement par habitant avec les deux types de matrices.

L'impact de la base sur le foncier non bâti sur les dépenses d'équipement est positif sur la période récente ce qui pourrait traduire une politique liée à la disponibilité foncière latente en vue de son urbanisation. Les territoires en tension dans le périmètre du Grand Paris peuvent être concernés.

Conclusion

Nous avons développé dans ce chapitre, un modèle de *Difference-in-Difference* spatial (SDID) pour estimer l'effet des CDT sur les dépenses d'équipement par habitant des communes avec prise en compte des interdépendances dans le cadre du projet Grand Paris. Nous avons modélisé ces effets d'interdépendance sous forme de dépendance spatiale avec un modèle SAR avec deux matrices de poids différentes : une matrice économique, basée sur les flux domicile travail entre les communes et une matrice géographique basée sur une distance seuil de 10 km.

Dans un premier temps, les résultats du diagnostic d'autocorrélation spatiale conduisent au rejet de l'hypothèse nulle d'absence d'interdépendances spatiales entre les communes. En conséquence, les estimations de l'effet des CDT avec les moindres carrés ordinaires sont biaisées et les conclusions sur l'effet réel des CDT sont également biaisées.

Les estimations du modèle SAR montrent, avec les deux matrices, que le paramètre d'interdépendance spatiale est significatif et négatif. Cela implique qu'il y a des externalités horizontales de « *complémentarité* » dans les dépenses d'équipement par habitant entre les communes selon le type de voisinage choisi. Lorsque les communes j voisines de i augmentent « en moyenne » leurs dépenses, alors la commune i diminue les siennes. Ce comportement des communes peut s'expliquer à la fois par un comportement stratégique et par l'appartenance de la commune à un EPCI.

Les résultats montrent que les effets estimés des CDT sur les dépenses d'équipement sont différents selon les périodes. Entre 2001-2008, l'effet pré-annonce du Grand Paris est non significatif, ce qui signifie qu'il n'y a pas eu de comportements d'anticipation des communes par rapport à leurs dépenses d'équipement.

En revanche, sur la période post-annonce (2008-2012), les résultats indiquent que l'annonce du projet semble avoir un impact positif et significatif (au seuil de 10%) sur les dépenses des communes CDT par rapport aux communes hors CDT.

Quant à la période globale (2001-2012), l'effet *ex ante* du projet est également positif et significatif (au seuil de 5 %) sur les dépenses d'équipement des communes. Cela veut dire que le projet Grand Paris semble avoir un impact positif sur les dépenses d'équipement des communes CDT.

Concernant les effets des interactions, ils sont présents sur la période post-annonce et sur la période globale. Ils sont significatifs et négatifs, ce qui veut dire que la commune diminue ses niveaux de dépenses lorsque les communes voisines augmentent les leurs. Ceci peut être interprété comme étant une complémentarité des dépenses d'équipement entre les communes.

Ce travail contribue d'une manière plus générale au développement de méthode d'analyse des projets d'aménagement et plus spécifiquement à celui du Grand Paris. Il développe une méthode d'évaluation *ex ante* du projet dont la définition, la construction et la mise en œuvre prennent du temps. Des effets d'anticipation sont alors susceptibles de se mettre en place et dont il faudra tenir compte dans l'évaluation *ex post*. Notre travail permet de dresser une situation de référence sur le projet du Grand Paris et qui a été appliqué aux dépenses publiques des communes avant la mise en œuvre effective du projet.

Le prolongement de ce travail consisterait à prendre en compte l'aspect dynamique de l'intégration des communes dans les CDT.

Annexe A: Profil historique du Grand Paris entre 2007 et 2010⁴⁶

- 7 Juillet 2010 : Présentation du décret instituant la Société du Grand Paris
- 16 Juin 2010 : Les collectivités de droite peuvent adhérer à Paris Métropole
- 3 Juin 2010 : Promulgation de la loi relative au Grand Paris**
- 27 Mai 2010 : La loi sur le Grand Paris adoptée par le Parlement
- 20 Mai 2010 : Les députés et sénateurs répartis en commission paritaire
- 15 au 16 Avril 2010 : Séminaire d'appel à initiatives Paris Métropole
- 7 Avril 2010 : 200 élus de gauche dénonce le projet du Grand Paris
- 6 Avril 2010 : Examen au Sénat de la loi sur le Grand Paris
- 21 Mars 2010 : Victoire de la Gauche aux élections régionales
- 25 Février 2010 : E.Balladur propose un référendum sur le Grand Paris
- Février 2010 : Création de l'Atelier International du Grand Paris
- 18 Décembre 2009 : Première étude territoriale du Grand Paris
- 17 Décembre 2009 : Premier forum Paris Métropole à la rencontre des citoyens de la métropole
- 15 Décembre 2009 : N. Sarkozy reçoit les 10 équipes d'architectes-urbanistes
- 1^{er} Décembre 2009 : L'Assemblée Nationale adopte le projet de loi sur le Grand Paris
- 29 Octobre 2009 : Constitution de l'Atelier international du Grand Paris
- 30 Septembre 2009 : Christian Blanc critique la version du projet de loi transmise au Conseil d'État
- 25 Septembre 2009 : Paris Métropole définit une position face au projet de loi du Grand Paris
- 7 Octobre 2009 : Le projet de loi sur le Grand Paris
- 1^{er} Octobre 2009 : Publication du rapport Carrez
- 9 Septembre 2009 : La Commission Nationale du Débat Public approuve l'organisation du débat sur Arc Express
- 29 Juillet 2009 : Proposition SNCF de projet de liaison par métro automatique
- 16 Juillet 2009 : "Faire du Grand Paris une véritable Métropole maritime" discours du président
- 9 Juillet 2009 : Saisine de la Commission Nationale du Débat par le STIF
- 8 Juillet 2009 : Adoption d'Arc Express au STIF
- 10 Juin 2009 : Lancement officielle de Paris Métropole
- 5 Juin 2009 : Rapport Carrez sur le financement des infrastructures de transport
- 30 Avril 2009 : Validation des statuts de Paris Métropole
- 29 Avril 2009 : Inauguration par N. Sarkozy de l'exposition "Le Grand Pari(s) de l'agglomération parisienne"
- 30 Mars 2009 : Séance de la Conférence Métropolitaine à Paris. Échanges sur le Grand Pari(s) avec les 10 équipes de la consultation internationale
- 5 Mars 2009 : Publication du rapport Balladur sur la réforme territoriale.
- 16 Janvier 2009 : Réunion entre les 10 équipes et des élus d'IDF
- 5 Novembre 2008 : Finalisation des statuts de Paris Métropole
- 10 Octobre 2008 : Plateau de Saclay (Christian Blanc présente les orientations qui guideront le développement du plateau de Saclay)
- 1^{er} Octobre : Élaboration des statuts de Paris Métropole
- 25 Septembre 2008 : Adoption du SDRIF de 2008
- 25 Juin 2008 : Assises de la métropole

⁴⁶ Historique du Grand Paris <http://www.ateliergrandparis.fr/grandparis/> , consulté 22/09/2015

- 4 Juin 2008 : lancement officiel de la consultation internationale sur le Grand Paris
- 8 Avril 2008 : Publication du rapport Dallier sur le Grand Paris
- 18 Mars 2008 : Création d'un secrétariat d'Etat pour le développement de la région (nomination de Christian Blanc comme secrétaire d'Etat)
- 27 Décembre 2007 : Consultation internationale du Grand Paris
- 15 Février 2007 : Le projet de SDRIF est adopté par le Conseil régional d'Île-de-France
- 25 Octobre 2007 : Création d'une commission "Scenarii pour la métropole: Paris-Ile-de-France demain"
- 17 Septembre 2007 : en inaugurant la Cité de l'architecture, Nicolas Sarkozy annonce son souhait d'un nouveau projet d'aménagement global du Grand Paris
- 26 Juin 2007** : Discours-programme de N.Sarkozy sur la région IDF : "Construisons enfin des transports en commun circulaires". Lors de l'inauguration d'une nouvelle aérogare à CDG et en présence de B. Delanoë, N. Sarkozy, Président de la République, prononce son discours-programme sur la région IDF et expose sa vision du Grand Paris

Chapitre 4 Hétérogénéité des interactions et évaluation des effets de traitement : Une approche par les effets de dépendance spatiale⁴⁷

⁴⁷ Ce chapitre correspond à une version sommairement modifiée de l'article Ba S., Baumont C. (2015), publié sous le même titre dans la Revue d'Économie Régionale et Urbaine 2015/1 (mai), p. 105-148.

Introduction

En préconisant à la fin de l'année 2013, d'amplifier les évaluations des politiques publiques concernant les entreprises, les domaines de l'éducation ou encore de la politique de la ville, la Cour des Comptes soulignait les économies de plusieurs milliards qui pourraient être réalisées notamment via une meilleure utilisation des fonds alloués à ces politiques⁴⁸. Les méthodes économétriques d'évaluation des politiques publiques et, parmi elles, celles des effets de traitement (Heckman et Vytlačil, 2007a; Rubin, 1974) que nous considérons dans cette contribution, sont concernées par cette recommandation. Idéalement, l'évaluation des effets de traitement consisterait à pouvoir isoler des résultats observés, ceux qui sont imputables à la politique et à qualifier ces résultats en fonction des différents rouages ou mécanismes qui les ont favorisés ou contraints. L'évaluation peut alors produire un ensemble de recommandations sur les mécanismes socio-économiques à l'œuvre lorsque des dispositifs sont mis en place en vue d'atteindre des objectifs de développement économique (création d'emplois, création ou croissance des entreprises...), de réduction des inégalités de revenu ou d'accès à l'emploi, à l'éducation ou au logement, de renouvellement urbain, de développement des zones urbaines sensibles...

Dans la réalité, la mise en œuvre d'une politique et ses effets sont le plus souvent marqués par des relations complexes impliquant de nombreux acteurs et qui vont impacter à la fois le processus d'assignation du traitement (sélection des bénéficiaires et des non-bénéficiaires) et les performances de la politique. La possibilité de mener des programmes expérimentaux constituerait une manière idéale d'isoler ces effets, de contrôler les conditions de sélection des individus et les interférences pouvant exister entre ceux qui sont bénéficiaires et leur entourage. Dans le cadre non expérimental, la littérature va de fait s'appuyer principalement sur deux hypothèses qui restreignent la prise en compte des effets engendrés par différentes formes d'interférences. Premièrement, l'hypothèse CIA (*Conditional Independence Assumption*, (Imbens et Wooldridge, 2009) postule l'indépendance conditionnelle entre le traitement, c'est-à-dire la sélection des bénéficiaires, et le résultat obtenu. La non validité de l'hypothèse CIA interroge l'évaluateur sur l'approche à mener « en amont » de l'évaluation des effets de la politique : il s'agit de s'assurer que l'on a suffisamment bien appréhendé les bénéficiaires de la politique, tant par leurs caractéristiques que par les raisons qui auraient pu amener certains individus à adhérer au dispositif ou au

⁴⁸ Modernisation de l'Action Publique (MAP), décembre 2012.

contraire à s'en détourner. Deuxièmement, l'hypothèse SUTVA -*Stable Unit Treatment Value Assumption* - (Rosenbaum et Rubin, 1983 ; Rubin, 1974) postule que le traitement d'un individu n'affecte que le résultat de cet individu et n'a donc aucun effet sur les résultats des autres individus. Cette hypothèse s'applique en outre, dans sa définition stricte, aux seules relations entre les individus bénéficiaires. La non validité de cette hypothèse interroge l'évaluateur sur la pertinence des résultats de l'évaluation, notamment parce que la présence d'interactions remet en cause la performance des estimateurs.⁴⁹

Partant du constat que les interactions entre les individus existent et font partie intégrante des comportements et résultats observés, il est également difficilement tenable de les ignorer. En effet, il résulte des interactions entre les individus, des effets supplémentaires, favorables ou contraignants, modifiant l'effet final via les externalités et les effets multiplicateurs. Les interactions sociales, présentées par Manski (1993), entre autres, soulignent l'influence des pairs, des normes sociales... qui sous-tendent des mécanismes d'imitation, de conformité... avec des effets de propagation ou de diffusion au sein des groupes. Le rôle des réseaux a été largement souligné comme support des interactions sociales (Calvo-Armengol et *al.*, 2009 ; Moffit, 2001 ; Topa, 2001). Par ailleurs, les interactions spatiales mettent en avant le rôle de la proximité géographique entre les individus comme vecteur favorisant ces interactions (Topa, 2001). Les interactions spatiales et sociales sont combinées, par exemple, dans le cadre de politiques sanitaires, de programmes de vaccination ou de prévention d'une épidémie : plus le nombre de personnes traitées sur la zone sera important et plus le risque sera réduit pour les personnes non vaccinées (effet de diffusion par la proximité), et la politique sera d'autant plus facilement acceptée que le nombre de personnes s'y soumettant est grand ou qu'il s'agit des personnes influentes (effets de diffusion par les pairs ou rôle des leaders). La présence d'externalités spatiales interroge en outre sur les modalités de l'évaluation. Rathelot et Sillard (2008), par exemple considèrent, dans le cas des politiques de Zones Franches Urbaines (ZFU), la situation des zones classées ZFU et celle des zones situées dans leur voisinage immédiat, à moins de 300 mètres. Mayer et *al.* (2012) proposent de contrôler au sein des communes qui disposent d'un quartier ZFU, l'attractivité des autres quartiers non ZFU, car le choix de localisation peut se faire à deux niveaux (commune et quartier) et la différenciation spatiale s'apprécier finement dans ces choix. Dans ces études, on peut s'attendre par exemple, à une augmentation de l'attractivité des zones

⁴⁹ Les situations rencontrées dans un cadre expérimental sont nuancées par la suite dans la Section 2.

ZFU du fait des dispositifs associés au détriment des zones non ZFU ou, à un effet contraire, si le classement en ZFU apparaît comme un signal négatif. Le développement des zones ZFU peut aussi conduire à un développement des zones non ZFU en cas de saturation des premières. Einio et Overman (2011) interrogent également la validité de l'évaluation de ces politiques zonées du fait des externalités géographiques. Yamazaki et Resosudarmo (2008), analysent la politique *Farmer Field School* en Indonésie sur l'adoption des nouvelles technologies agricoles en insistant sur la présence d'interactions sociales et géographiques : du fait des interactions sociales inhérentes à l'appartenance à une même communauté, les personnes d'un même village sont considérées comme « exposées » et peuvent donc adopter la politique alors qu'entre villages différents, la diffusion de la politique n'est plus assurée. Cet effet de transmission entre les individus formés et leur voisins au sein du village a permis un accroissement de la productivité agricole chez les non participants à travers la transmission de nouvelles compétences, effet non mis en évidence dans une étude préalable (Feder et al., 2004). Avec une autre approche des externalités spatiales, une étude réalisée par Affuso (2010) porte sur l'efficacité des agriculteurs participant à un programme de vulgarisation agricole en Tanzanie. L'auteur estime que les effets de contagion spatiale à la frontière pourraient générer un gain de 5,7 % pour les non participants. L'étude souligne par ailleurs que le rendement moyen technique peut doubler si les processus spatiaux sont inclus dans le modèle.

L'objet de notre recherche est de contribuer à cette littérature sur la prise en compte des effets des interdépendances dans l'évaluation des politiques publiques, dans un cadre non expérimental, en considérant les interactions entre les individus et leur dimension spatiale. Plus précisément, nous proposons trois extensions au modèle d'évaluation en deux étapes de Heckman (Heckman et al., 1998 ; Heckman, 1979). Premièrement, nous considérons un schéma général pour les interactions qui peuvent concerner tous les individus, qu'ils soient bénéficiaires ou non, et nous proposons de les caractériser par les matrices de voisinage spatial et de les modéliser par les spécifications économétriques spatiales. Deuxièmement, nous proposons, dans ce cadre, d'étendre le processus de sélection aux caractéristiques inobservables des individus. En effet, il est généralement supposé que la sélection des individus est traitée par un ensemble de déterminants observables, car pour être bénéficiaire, il faut remplir un certain nombre de critères. Pourtant, certains individus éligibles peuvent ne pas participer pour différents motifs de renoncement liés à des facteurs d'anticipation du résultat potentiel ou par refus de la stigmatisation associée au « traitement ». Classiquement,

ces formes d'auto-sélection peuvent engendrer un biais de sélection : la population sur laquelle le traitement est appliqué ne représente pas la « vraie population ». En outre, ces individus, compte tenu de leurs caractéristiques, risquent de se retrouver à tort comme des individus de contrôle dont l'évolution est différente de celle des individus traités. Plus globalement cependant, la décision de participer ou non peut être influencée par un ensemble de facteurs inobservables comme les participations aux réunions de sensibilisation, les opinions ou mots d'ordre de leaders ... associés aux voisinages sociaux et géographiques des individus. Selon Manski (1993), les interactions sous-jacentes sont alors plutôt attribuables à des facteurs institutionnels qu'il nomme « effets corrélés ». Nous proposons alors de modéliser ces interdépendances par un processus autorégressif spatial sur les erreurs dans l'équation de sélection et ainsi de corriger le biais de sélection sous-jacent. Troisièmement, nous prenons en compte la diversité des interactions dans l'estimation des équations de résultats. Un individu selon son statut (bénéficiaire ou non) peut avoir un voisinage hybride ou hétérogène composé d'individus traités, ou d'individus non traités, ou les deux à la fois. Nous souhaitons estimer si l'hétérogénéité des voisinages a des effets différents. Nous modélisons les équations de résultats, pour les bénéficiaires et le groupe de contrôle, sous la forme d'un modèle spatial autorégressif de telle sorte que le résultat d'un individu dépend du résultat des autres individus selon leur statut. Cette proposition permet de tenir compte des effets d'interaction endogènes au sens de Manski (1993) et elle étend, au cas des interactions entre individus bénéficiaires et non bénéficiaires, le cadre proposé par Flores-lagunes et Schnier (2012) . Elle constitue sur le plan méthodologique un élargissement du cadre d'invalidité des hypothèses SUTVA et PIA au cas des interactions intergroupes.

Le développement de ces propositions est organisé de la manière suivante. Dans la deuxième section, nous proposons de revenir sur les interrogations soulevées par la prise en compte des interactions dans le modèle d'évaluation de l'effet causal pour positionner les marges de progression de notre contribution. La section 3 est consacrée à la présentation du modèle économétrique spécifié dans le cadre spatial et à ses propriétés. Dans la section 4, nous proposons une application empirique de notre modèle. Nous nous intéressons au cas de la politique de revitalisation des zones rurales (ZRR) et comparons les résultats du modèle standard avec ceux obtenus dans notre modèle spatial. Nous montrons sur cette application l'hétérogénéité des interactions : la création d'établissements dans les ZRR dépend de la création d'établissements dans les communes ZRR voisines et de la création d'établissements dans les communes voisines non classées en ZRR. Notre application montre également que si

le modèle standard conclut à un effet de la politique spécifique aux communes ZRR, cet effet disparaît lorsque les interactions entre communes voisines sont prises en compte. La section 5 met en débat et perspectives de recherches futures, les limites et extensions au modèle d'évaluation des politiques publiques dans un contexte d'interactions spatiales et sociales.

4.1. Interactions, interdépendances et évaluation des politiques publiques

Nous nous situons dans un cadre non expérimental et supposons que l'affectation au traitement n'est pas obligatoire. Les notations et principes du modèle d'évaluation de l'effet causal en deux étapes (Heckman et *al.*, 1998 ; Heckman, 1979) sont fixés dans un premier temps, avant de préciser le rôle et les effets de la prise en compte des interactions dans ce modèle.

4.1.1. Le cadre conceptuel du modèle d'évaluation

4.1.1.1. Le modèle d'évaluation en deux étapes

Nous considérons le cas des modèles de choix discret dans lesquels on est en présence de deux types d'individus : les bénéficiaires du programme appelés « groupe des traités » et les non bénéficiaires. Ces derniers constitueront les individus du groupe de contrôle. On suppose que chaque individu i a deux résultats potentiels, selon qu'il est bénéficiaire ou non du traitement. La différence entre les deux résultats Δ_i représente l'effet propre ou causal de la politique sur l'individu i :

$$\Delta_i = y_{i1} - y_{i0} \quad (4-1)$$

y_{i1} est la variable de résultat potentiel si l'individu i a reçu le traitement et y_{i0} s'il n'a pas reçu le traitement. Dans la mesure où l'on ne peut pas observer simultanément les deux résultats pour un même individu, on cherche à estimer une situation contrefactuelle indiquant ce qui aurait été observé chez l'individu traité si la politique n'avait pas été mise en place. On peut alors construire une situation contrefactuelle au mieux satisfaisante, dans laquelle le groupe de contrôle serait constitué d'individus semblables aux bénéficiaires mais qui n'ont pas été sélectionnés dans le programme.

La construction du contrefactuel repose formellement sur la méthode suivante. Soit T^* une variable latente qui reflète la propension d'un individu à être traité ($T^*= 1$) ou non traité ($T^*= 0$). Cette propension est fonction d'un ensemble de déterminants regroupés dans un vecteur de variables explicatives Z . Le modèle structurel de la variable latente s'écrit de la manière suivante :

$$T^* = Z\alpha + u \quad (4-2)$$

α est un vecteur de paramètres à estimer et le résidu u suit une distribution normale.

La variable latente T^* étant non observée, on a une variable binaire observée T liée à T^* à travers la relation suivante :

$$T_{ik} = \begin{cases} 1, & \text{si } T_{ik}^* > 0 \\ 0, & \text{si } T_{ik}^* \leq 0 \end{cases} \quad (4-3)$$

Par la suite, l'indice k prendra la valeur 1 pour les individus traités et la valeur 0 pour les non bénéficiaires.

L'estimation de l'équation de sélection apporte pour l'évaluateur un ensemble d'indications sur le processus de sélection en fonction des caractéristiques des individus. Cette estimation permet par ailleurs de considérer une caractéristique assez fréquente du processus de sélection lui-même : le biais de sélection. Celui-ci est présent lorsque l'affectation des individus à des groupes de traitement et de contrôle n'est pas aléatoire, quand bien même la politique publique s'adresse à des bénéficiaires bien identifiés, en fonction de caractéristiques socio-économiques précises : âge, statut social, niveau d'étude, niveau de revenu, secteur d'activité, taille d'entreprise, lieu d'habitat... Pourtant, les individus qui remplissent les conditions observables et bénéficient effectivement du traitement peuvent participer en définitive sur la base de caractéristiques inobservables : les bénéficiaires peuvent par exemple avoir reçu une meilleure information sur la politique. De même, certains individus éligibles ont pu se soustraire volontairement du programme sur la base de ces informations inobservables. Dans ce cas, on parlera de biais d'auto-sélection (Barnow et *al.*, 1980 ; Goldberger, 2008). Si l'existence d'un biais de sélection est avérée, cela implique que les estimateurs de l'effet causal sont surestimés ou sous-estimés et qu'il convient de le corriger (Brodsky et *al.*, 2007). A cet effet, l'estimation de l'équation de sélection permet de calculer

l’Inverse du Ratio de Mills (IRM) qui sera intégré dans les équations de résultats à la seconde étape.

Pour chaque catégorie d’individus, les variables de résultats latents sont supposées être chacune engendrée par le modèle de régression linéaire suivant :

$$Y_k = X\beta_k + \varepsilon_k \quad \text{avec } k = \{0,1\} \quad (4-4)$$

X est une matrice de variables explicatives, β_k est un vecteur de paramètres associés et ε_k est le terme d’erreur du modèle suivant les propriétés usuelles.

Finalement, l’évaluation de l’effet causal repose sur l’estimation de l’impact de la politique sur les individus traités et non traités et dans la littérature, on s’intéresse généralement à deux types d’effets.

L’effet moyen de traitement sur les traités, uniquement, noté Δ^{ATT} :

$$\Delta^{ATT}(X, Z) = E(Y_1 - Y_0 | T = 1, X, Z) \quad (4-5)$$

L’effet moyen de traitement sur la population totale, c’est-à-dire l’effet global de la politique si elle était étendue à l’ensemble de la population, noté Δ^{ATE} :

$$\Delta^{ATE}(X, Z) = E(Y_1 - Y_0 | X, Z) \quad (4-6)$$

4.1.1.2. Le contrôle des interdépendances

Généralement, on stipule l’absence d’interdépendances au sein du modèle à deux niveaux. Premièrement, l’hypothèse CIA⁵⁰ – *Conditional Independence Assumption* – (Imbens et Wooldridge, 2009 ; Angrist et Pischke, 2009; Lechner, 2001) suppose qu’il n’y a pas d’autres caractéristiques inobservables qui affectent à la fois le processus d’affectation au traitement et le résultat. Un cas particulier serait alors de dire que le résultat de la politique n’est pas conditionnel au processus de sélection. Deuxièmement, l’hypothèse SUTVA⁵¹ – *Stable Unit Treatment Value Assumption* – (Rosenbaum et Rubin, 1983)–postule que le traitement d’un individu affecte uniquement le résultat de cet individu et n’a aucun effet sur le

⁵⁰ Ou encore *Unconfoundedness Assumption* pour Imbens et Wooldridge (2009) .

⁵¹ On la retrouve également sous l’appellation *Individualistic Treatment Response* (ITR) pour Manski (1993).

résultat des autres individus. Cette hypothèse est posée au sein du groupe des individus bénéficiaires mais on pourrait la décliner en toutes généralités au sein du groupe des non bénéficiaires et entre les individus de groupes différents. C'est notamment sur cette extension que repose notre proposition. De même, si cette hypothèse est spécifiée au niveau des équations de résultats, il apparaît naturel de s'interroger aussi sur sa pertinence au niveau de l'équation de sélection. Dans ce cas, l'existence d'interactions entre les individus signifie que les décisions des uns de participer ou non au programme ne sont pas indépendantes de ce que les autres font (Brock et Durlauf, 2001) . Cette perspective est également discutée dans notre proposition.

4.1.2. Les interdépendances sociales et géographiques

Ainsi, le domaine de validité du modèle d'évaluation de l'effet causal de Rubin est subordonné à l'absence d'interférence entre les individus et entre le traitement et le résultat. Ce contexte très particulier peut être satisfait dans le cadre de programmes expérimentaux.⁵² En dehors de ces situations contrôlées, les cas d'interdépendance sont au contraire nombreux. Les mécanismes qui structurent ces interdépendances sont bien connus des économistes et désignent les situations d'interactions stratégiques, les externalités ou les spillovers. Les développements sur les interactions spatiales et sur les interactions sociales sont particulièrement concernés. En effet, dans le contexte des politiques publiques qui sont « territorialisées » du fait d'espaces ciblés, ou « localisées » lorsque ce sont les individus qui sont ciblés, la nature des interdépendances sera plutôt d'ordre géographique dans le premier cas et d'ordre social dans le second cas. Ces deux champs peuvent être appréhendés par un ensemble d'outils communs.

La dimension géographique des données est un marqueur désormais bien connu de la présence d'interdépendances qui se manifeste par une organisation spatiale non aléatoire des phénomènes : « *many econometric textbooks are aware of the problem of non independence arising when dealing with spatially collected data* » (Arbia, 2006, p. 90). On distingue notamment la dimension absolue et la dimension relative (Abreu et al., 2005). La dimension *absolue* est inhérente à la localisation : toute politique publique territorialisée ou zonée correspond à ce cas, par définition, car elle implique un périmètre d'intervention. Les espaces

⁵² Ce qui ne signifie pas bien entendu que les programmes expérimentaux excluent la présence d'interdépendances. Le lecteur pourra se reporter à l'article récent de BAIRDS et al, 2014, pour une présentation de méthodes de désignation des ensembles d'individus et d'évaluation des effets de spillovers.

sont différenciés en fonction de leur localisation et des variables indicatrices peuvent être utilisées pour apprécier les effets fixes géographiques potentiels. La dimension absolue peut de ce fait être assimilée à l'hétérogénéité spatiale. La dimension géographique est aussi *relative* lorsqu'elle fait référence à l'organisation spatiale, c'est-à-dire à la manière dont les territoires se situent les uns par rapport aux autres, tissent entre eux des interactions de différentes natures : interactions marchandes et non marchandes via l'échange de biens et services, les déplacements de personnes, les échanges d'informations ou de savoirs, les relations sociales et culturelles, les réseaux d'affaires, les réseaux informels... L'organisation spatiale peut alors être décrite par des schémas géographiques d'interconnexions qui permettent de repérer les relations entre eux et d'apprécier la force de ces relations. Dans ce cas, on fait référence à l'autocorrélation spatiale. Par extension, la distribution spatiale d'un phénomène associe souvent l'hétérogénéité spatiale et l'autocorrélation spatiale sous la forme de schémas de disparités spatiales : par exemple on aura une tendance au regroupement des quartiers riches dans certaines parties d'une agglomération et des quartiers pauvres dans d'autres parties de l'agglomération. Enfin, l'organisation spatiale des données n'est pas seulement un schéma que l'on observe, mais c'est également un facteur explicatif des phénomènes observés. Par exemple, la concentration spatiale d'élèves en situation d'échec scolaire peut être un facteur aggravant l'échec scolaire. Si un territoire se situe à proximité de zones économiques en croissance, cela peut être un facteur favorisant son dynamisme économique. La concentration d'industries spécialisées peut être un facteur favorisant l'innovation. La pollution est plus forte près des zones de concentration des véhicules. Dans tous ces exemples, l'organisation spatiale des données est porteuse d'externalités spatiales, positives ou négatives, ou d'effets de débordement géographique.

Les interactions sociales désignent les interdépendances entre les individus et ciblent principalement le fait que le comportement d'un individu peut être *directement* influencé par les comportements des autres individus, c'est-à-dire en se situant en dehors du marché. On peut alors « observer » que les actions d'un individu ne sont pas indépendantes des actions des autres individus sous l'impulsion, pour l'approche sociologique, d'effets de voisinage, d'effets de pairs ou d'imitation, des normes sociales ..., ou du fait d'externalités pour les économistes. Il est devenu usuel d'utiliser la typologie proposée par Manski (1993) pour définir les trois formes, non exclusives, expliquant pourquoi un individu peut se comporter comme les autres individus. Premièrement, les interactions endogènes s'appliquent lorsque le comportement d'un individu dépend directement du comportement des autres individus.

Deuxièmement, les interactions exogènes ou contextuelles s'appliquent lorsque la propension d'un individu à adopter un certain comportement dépend des caractéristiques, exogènes, de son environnement. Enfin, les effets corrélés s'appliquent lorsque la propension d'un individu à se comporter comme les autres individus provient de caractéristiques individuelles ou institutionnelles similaires. Cette terminologie est, selon Manski (1993), plus précisément applicable au sein d'un groupe d'individus « similaires » ce qui dans notre cas fait référence au groupe des bénéficiaires ou à celui des non bénéficiaires. Comme pour les interactions spatiales, l'entourage des individus ou les réseaux avec lesquels ils sont en contact ne peuvent a priori pas être négligés ou considérés comme neutres.

Conceptuellement, les interactions spatiales ou sociales partagent des points communs puisque la modélisation des interactions endogènes, exogènes et corrélées peut être réalisée via les spécifications d'économétrie spatiale (Soetevent, 2006). L'existence de ces interactions, dans certains cas, remet en cause l'estimation des équations par les méthodes usuelles.

4.1.3. Interactions et évaluation des politiques publiques

Considérer la question des interactions dans l'évaluation des politiques publiques revient à intégrer conceptuellement le fait que les individus ne sont pas isolés les uns des autres et que dans bien des situations les objectifs qu'ils peuvent atteindre sont non seulement dépendants de leur choix mais aussi de celui des autres. La prise en compte des interdépendances se pose à trois niveaux : qui interagit avec qui, quand et sous quelles formes ? Au sein du modèle décrit précédemment, cette conceptualisation intervient au niveau de l'équation de sélection et au niveau de l'évaluation de l'effet causal, sauf dans le cadre de conditions expérimentales qui s'apparentent aux situations les plus appropriées pour éviter les distorsions sur la sélection et imposer l'absence d'interactions entre les individus. Un cadre expérimental, ayant contrôlé en amont les processus de sélection, peut, cependant, vouloir apprécier l'impact des interactions sur les bénéficiaires du programme et dans ce cas, l'hypothèse SUTVA n'est pas respectée.

La modélisation du processus de sélection permet d'identifier les déterminants de l'affectation d'un individu dans le groupe des bénéficiaires ou des non bénéficiaires. Ces derniers constituent, sous certaines conditions, le groupe de contrôle utile à l'estimation de l'effet causal. Si le problème du biais de sélection a fait l'objet de nombreux travaux avec

différentes méthodes pour le traiter (Brodaty et *al.*, 2007 ; Maddala, 1983), la question des interactions reste relativement peu développée alors qu'elle amène également à s'interroger sur la validité de l'hypothèse CIA. Dans le cas des politiques de clusters, par exemple, le processus de sélection est défini par les pouvoirs publics, mais il peut en définitive amener les entreprises, familières de ces dispositifs, à s'auto-sélectionner sur la base d'interactions entre elles et dans la perspective des résultats escomptés.⁵³ Dans la mesure où un grand nombre de déterminants peuvent être impliqués dans les interdépendances au sein du processus de sélection, une attention particulière doit aussi être apportée aux caractéristiques inobservables. Ceci rejoint une modélisation d'un processus spatial autorégressif du terme d'erreurs en référence aux interactions corrélées Manski (1993).

Le problème de l'endogénéité du traitement peut être traité par d'autres estimateurs, variables instrumentales ou discontinuité de la régression (RDD) par exemple. Cependant, la présence d'interactions entre les individus soulève plusieurs problèmes. D'une part, dans le cas non spatial, l'identification des bons instruments a déjà été soulignée dans le cas des équations de sélection avec inobservables (Brodaty et *al.*, 2007) et serait aussi problématique dans le cas spatial. Selon (Givord, 2010), les propriétés des estimateurs ne sont pas satisfaisantes ce qui amène des questions quant à la fiabilité des estimations. D'autre part, l'estimateur RDD, mis en œuvre notamment par Lorenceau (2009) permet de considérer le problème du biais de sélection lorsque des « *possibilités de détournement des seuils* » existent. Par exemple, certaines entreprises peuvent se délocaliser des quartiers non bénéficiaires d'un dispositif vers des quartiers bénéficiaires (effet d'aubaine) ou inversement (effet stigmatisant). L'approche est intéressante dans le cas de facteurs inobservables notamment. Pour autant, l'estimateur fournit une évaluation du dispositif pour les individus proches du critère d'éligibilité et il s'agit d'un effet local ce qui pose le problème de la validité externe des résultats obtenus. Dans notre approche, la prise en compte des interactions entraîne un effet global, car nous estimons les impacts induits sur l'ensemble des individus, qu'ils soient proches du seuil ou non.⁵⁴

⁵³ Notons que l'obligation ou le libre choix de participer ou non à la politique ne sont pas en cause ici, car même en cas d'obligation, comme par exemple dans le cas de l'article 55 de la loi SRU sur le logement social, un certain nombre d'acteurs peuvent être influencés par leurs pairs dans leur choix de se « soustraire » à l'obligation et « d'accepter » de payer les amendes.

⁵⁴ Le lecteur peut se reporter à l'équation (4-8) pour visualiser cette propriété.

Ces interactions présentes en amont et contrôlées, ne sont pas mécaniquement une cause d'invalidation de l'hypothèse SUTVA. Elles peuvent constituer en revanche un signal quant aux sources d'interactions possibles au sein de la population et qui pourraient se reproduire dans l'application de la politique, c'est-à-dire après la sélection. Par exemple, les analyses sur le rôle des effets de voisinage sur la réussite scolaire (Overman, 2002) préconisent de ne pas tenir nécessairement comme équivalent - ni radicalement différent - ce qui se passe au sein de la famille (ou du quartier de résidence, dans des voisinages réduits) et qui pourrait plus influencer la participation, et ce qui se passe dans les lieux d'application de la politique (les classes, les écoles, les voisinages plus larges...) pour les effets sur les résultats.

Au niveau des équations de résultats, par définition, la présence d'interactions invalide l'hypothèse SUTVA. Plusieurs travaux ont traité cette question et nous en soulignons plus précisément deux aspects : la dimension des interactions et leur modélisation.

Par dimension des interactions, nous désignons l'ensemble des individus qui interagissent entre eux. La plupart des travaux postulent des hypothèses qui limitent les effets des interactions aux individus du même groupe de référence comme le stipule l'hypothèse *Control Treatment Response* (CTR) (Manski, 2013) ou *Partial Interference Assumption* (PIA) (Hudgens et Halloran, 2008 ; Rosenbaum, 2007). Si dans un cadre expérimental, on peut contrôler cette situation, considérer que les individus traités n'interagissent pas avec les individus non traités semble difficilement réaliste ailleurs. Cela pose en tout cas un cadre restrictif qui ne permet pas de prendre en compte l'ensemble des interactions et par conséquent l'hétérogénéité des voisinages. En effet, lorsqu'un territoire est éligible à un dispositif, les Zones Franches Urbaines en France par exemple, il répond à des critères et peut être entouré de territoires non éligibles, car ne répondant pas à ces critères. La fixation de critères crée donc en soi une hétérogénéité entre les individus. De même, un élève bénéficiaire d'une politique éducative peut avoir dans son réseau ou voisinage à la fois des élèves bénéficiaires et des élèves non-bénéficiaires. Les voisinages sont donc hétérogènes. Par ailleurs, cette hypothèse peut être questionnée lorsque la définition d'un groupe de contrôle supposera d'identifier comme individus de référence, ceux qui ne sont pas bénéficiaires, mais qui présentent des caractéristiques proches. Les méthodes de régression par discontinuité utilisent notamment les seuils des critères d'éligibilité pour cela. L'hétérogénéité créée entre

les groupes du fait de la sélection peut difficilement être ignorée dans le cas où elle soutient de *facto* l'évaluation des effets. On peut revenir à ce propos sur le processus de sélection : les stratégies des acteurs peuvent même être induites par leur voisinage lorsqu'il est préférable de ne pas souscrire à la politique en espérant bénéficier des effets apportés par un voisinage de bénéficiaires. Un autre aspect associé à la dimension des interactions concerne la manière dont les interactions sociales sont structurées dans un groupe. Manski (2013) souligne par exemple qu'il existe des hiérarchies dans le voisinage notamment des leaders et des suiveurs qui mettent en exergue le fait que les interactions sont unidirectionnelles ou asymétriques.

La modélisation des interactions sous forme de dépendance spatiale dans les spécifications économétriques est recensée (Soetevent, 2006) et appliquée dans un ensemble de recherches qui implicitement se rallient à la proximité spatiale comme bon descripteur des potentiels d'interactions entre les individus. Dans ces recherches, la dimension des interactions reste spécifiée de manière restrictive sur le périmètre des interactions et le statut des individus. Par exemple, (Yamazki et Resosudarmo, 2007) ne retiennent que les interactions entre les individus d'un même groupe : les villageois peuvent être bénéficiaires ou non mais ils n'interagissent pas avec les individus d'un autre village. (Mcintosh, 2008) pose une hypothèse d'absence de spillovers « *No Spillovers Assumption* » qui implique aucun impact du traitement des personnes sélectionnées vers les personnes non sélectionnées au sein d'une même région. (Einio et Overman, 2011) posent l'hypothèse que le « traitement n'affecte pas les zones non traitées proches (...) ».

Dans la mesure où des effets multiplicateurs sont attachés aux interactions (Hartmann et al., 2008), les négliger totalement ou partiellement induit une mauvaise estimation des effets d'un programme et par conséquent une mauvaise allocation des ressources. Les études sous-estiment de ce fait les vrais bénéfices d'un programme si elles ne considèrent pas que les individus bénéficiaires peuvent transmettre leurs acquis à leurs voisins non participants ou si l'on néglige l'influence possible que peuvent exercer les voisins, selon leur statut, sur les prises de décision et les comportements d'un individu.

4.2. Modèle d'évaluation avec interactions spatiales

Pour estimer l'effet causal du traitement en présence d'interactions entre les individus, nous utilisons le modèle d'évaluation en deux étapes et estimons une équation de sélection et

les équations de résultats à l'aide de spécifications spatiales. Dans la première étape, nous faisons porter l'effet des interactions sur les caractéristiques inobservables de la population et estimons l'équation de sélection à l'aide d'un modèle *Probit spatial*⁵⁵ avec autocorrélation spatiale des erreurs. Dans la deuxième étape, nous estimons l'équation de résultat pour les individus traités et le groupe de contrôle en intégrant l'Inverse de Ratio de Mills spatialement ajusté comme variable explicative supplémentaire pour contrôler le biais de sélection (Flores-lagunes et Schnier, 2012). A la différence de Flores-lagunes et Schnier (2012), les équations de résultats sont modélisées pour tenir compte de l'hétérogénéité des voisinages : inter-groupes et intra-groupes. Ensuite, nous calculons les effets moyens de traitement intégrant ces interactions spécifiques.

4.2.1. L'équation de sélection spatialement ajustée

Dans le cas du processus de sélection, la présence d'effets spatiaux signifie que la probabilité qu'un individu participe ou non à un programme dépend non seulement de ses caractéristiques mais également de celles de son voisinage, dont une grande partie relève de l'inobservable. Si nous étudions, par exemple, les bénéficiaires de programmes pour le développement de l'agrotourisme, les individus localisés en des lieux bénéficiant d'aménités naturelles valorisables sur le plan touristique seront plus à même d'en bénéficier si leurs voisins en ont déjà fait l'expérience, comme leurs voisins pourraient à leur tour les imiter. Dans l'analyse des affectations d'usage des sols, la dépendance spatiale est également remarquée : l'usage agricole ou urbain des terres est lié à la situation des parcelles vis-à-vis des zones rurales ou urbaines et de l'usage en cours dans les parcelles voisines (McMillen, 1995). Dans le cadre du programme Farmer Field Schools « *to Fight Poverty and Foster Innovation in East Africa* », certaines personnes ont participé à ce programme parce que l'un de leurs voisins avait participé au programme dans le passé ou bien parce qu'elles habitent près du centre de formation. Ces éléments ont influencé de manière significative l'affectation au traitement en dehors des caractéristiques observables (Yamazaki et Resosudarmo, 2008). Case (1992) montre, dans le cadre d'une étude portant sur l'adoption de nouvelles technologies agricoles en Indonésie, que les leaders naturels peuvent influencer les futurs adoptants alors que l'influence de ces derniers est très faible sur les leaders naturels.

⁵⁵ On peut consulter Fleming (2004) pour une présentation des différentes contributions sur ce sujet.

Formellement, si l'on considère le processus de sélection précédemment décrit par les Equations (2) et (3), le modèle structurel de la variable latente T^* du *Probit spatial* s'écrit de la manière suivante :

$$T^* = Z\alpha + u \quad \text{avec } u = \lambda Wu + v \quad (4-7)$$

où Z désigne un ensemble de variables explicatives, où les erreurs suivent un processus autorégressif spatial et où v est un terme d'erreurs homoscédastiques et indépendantes. La variable latente T^* est donnée par l'équation (4-3).

Si $(I - \lambda W)$ est non singulière, on peut récrire à nouveau l'équation (4-7) sous la forme suivante :

$$T^* = Z\alpha + (I - \lambda W)^{-1}v \quad (4-8)$$

Le paramètre spatial λ permet d'estimer l'impact des facteurs inobservables sur la décision de participer ou non au programme. Les facteurs observables sont modélisés par les variables exogènes Z , notamment les critères d'éligibilité à la politique. Pour autant, un ensemble de déterminants sont inobservés et nous considérons que les interactions entre les individus en sont un élément structurant.

Pour une population de n individus, la modélisation des interactions est assurée par la matrice de pondération W de dimension $(n \times n)$ et de terme général w_{ij} . Par convention les termes diagonaux sont nuls. W permet de représenter la structure du voisinage et les valeurs des termes w_{ij} indiquent quels individus sont considérés comme interagissant entre eux et avec quelles intensités. Dans un contexte purement géographique, on peut considérer par exemple que les interactions se déroulent dans un rayon d'un kilomètre et que plus les individus sont près les uns des autres, plus ils interagissent. On peut également stipuler que les interactions sont fondées sur l'appartenance à une structure commune (une cité, une école, un club ou une association...). Un ensemble de déterminants non observables sont donc à l'origine des interactions possibles entre les individus et de leur ciblage potentiel en tant que personne effectivement bénéficiaire ou non de la politique. De ce fait, les évaluations

resteront, bien sûr, contingentes au choix de la matrice d'interactions qui est considérée comme exogène (Anselin, 1988).⁵⁶

Si $\hat{\lambda} = 0$ on retrouve un modèle Probit standard avec absence d'interactions entre les individus dans leurs choix. Si $\hat{\lambda} \neq 0$, on est en présence d'interdépendances entre individus. La structure autorégressive sur les erreurs permet également de considérer un effet de multiplicateur, via le terme $(I - \lambda W)^{-1}$ et indiquant, si $\lambda \neq 0$, qu'un choc aléatoire chez un individu affecte non seulement sa probabilité de participer mais aussi celle des autres individus de son voisinage de même que cet individu sera impacté par les chocs émanant du voisinage.

Formellement, la structure de la matrice de variance-covariance n'est plus homoscedastique et la fonction de vraisemblance s'écrit de la manière suivante (LeSage et Pace, 2009) :

$$L(T^*, W|\lambda, \alpha, \sigma^2) = \frac{1}{2\pi\sigma^{2(n/2)}} |I_n - \lambda W| e^{-\frac{1}{2\sigma^2}(v'v)} \quad (4-9)$$

avec

$$v = (I_n - \lambda W)(T^* - Z\alpha)$$

Soit pour la fonction de log-vraisemblance :

$$\ln L = \ln |I_n - \lambda W| - \frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{(T^* - Z\alpha)'[(I_n - \lambda W)'(I_n - \lambda W)]^{-1}(T^* - Z\alpha)}{2} \quad (4-10)$$

On obtient alors des estimateurs des paramètres α (équation (4-11)) et σ^2 (équation (4-12)) :

$$\hat{\alpha} = [Z'(I_n - \lambda W)'(I_n - \lambda W)Z]^{-1}Z'(I_n - \lambda W)'(I_n - \lambda W)T^* \quad (4-11)$$

$$\hat{\sigma}^2 = \frac{[(T^* - Z\hat{\alpha})]'(I_n - \lambda W)'(I_n - \lambda W)(T^* - Z\hat{\alpha})}{n} \quad (4-12)$$

En remplaçant les paramètres estimés $\hat{\alpha}$ et $\hat{\sigma}^2$ dans l'équation (4-10), on obtient la fonction de log-vraisemblance concentrée (équation (4-13)) :

⁵⁶ Ces approches n'ont pas pour objectif de proposer une méthode d'identification de la vraie structure d'interactions, c'est-à-dire celle qui correspondrait au véritable processus générateur des données. La structure des interdépendances n'est donc pas pleinement testable, mais peut être discutée en proposant différentes matrices W.

$$\ln L_c = C - \frac{n}{2} \ln \left(\frac{(T^* - Z\hat{\alpha})' [(I_n - \lambda W)' (I_n - \lambda W)]^{-1} (T^* - Z\hat{\alpha})}{n} \right) + |I_n - \lambda W| \quad (4-13)$$

$$\text{Avec } C = -\frac{n}{2} \ln(2\pi) - \frac{n}{2}$$

Plusieurs auteurs soulignent que l'estimation de l'équation (4-13) par la méthode du maximum de vraisemblance traditionnel n'est pas efficace et peu convergente (Billé et Arbia, 2013 ; Fleming, 2004). Plusieurs autres méthodes sont proposées : la méthode *Expectation-Maximization* (EM) utilisée par McMillen (1992, 1995), les méthodes d'estimation de types GMM de Pinkse et Slade (1998) ou encore les méthodes d'estimations bayésiennes développées par LeSage et Pace (2009). Suivant les analyses, différentes performances sont associées à ces méthodes, et les méthodes EM ou GMM présentent un ensemble de limites techniques ou économétriques. La validité de la méthode EM requiert de grands échantillons (LeSage et Pace, 2009 ; le Gallo, 2002 ; Billé et Arbia, 2013) et la matrice d'information ne peut pas être déterminée analytiquement (Le Gallo, 2002 ; Smirnov, 2010a). L'estimateur de Pinkse et Slade (1998), utilisé par Flores-lagunes et Schnier (2012), est une méthode d'estimation pour contrôler l'hétéroscédasticité mais elle nécessite de trouver les bons instruments et elle n'utilise pas la totalité de l'information disponible (Billé et Arbia, 2013). Les méthodes d'estimation bayésienne, en revanche, permettent de contrôler l'hétéroscédasticité et d'utiliser les informations des éléments non diagonaux de la matrice variance-covariance. Elles fournissent comme le souligne Le Gallo (2002) les mêmes estimations que celles données par la méthode du maximum de vraisemblance et sont robustes pour les petits échantillons.

4.2.2. Les équations de résultats spatialement ajustées

Les effets de la politique sont estimés de manière à pouvoir comparer les résultats obtenus par les individus qui ont bénéficié de cette politique (en nombre n_1) par rapport à ceux qui n'en ont pas bénéficié (en nombre n_0). Au total $n = n_1 + n_0$ individus sont observés. Par la suite, les individus sont ordonnés de telle sorte que les individus traités occupent les n_1 premières places et les individus non traités, les n_0 rangs suivants.

Nous considérons que les hypothèses SUTVA et PIA ne sont pas respectées du fait d'interactions au sein des groupes et entre les groupes et nous postulons des interactions endogènes au sens de Manski (2013). Les équations de résultats sont donc spécifiées sous la forme d'un modèle SAR. Cette modélisation permet de décrire les effets de diffusion au sein de la communauté et les effets multiplicateurs de la politique en termes de résultats, du fait du processus autorégressif sur les variables d'intérêts (équation (4-18)).

Considérons d'une manière générale, l'équation de résultats suivante :

$$Y = \rho_{k1} W_k^1 Y + \rho_{k0} W_k^0 Y + X\beta_k + \varepsilon_k \quad (4-14)$$

où Y représente la variable dépendante étudiée qui indique la performance de la politique (taux de réussite scolaire, taux d'insertion professionnelle, nombre de logements sociaux construits...). Si $k = 1$ on obtient l'équation de résultat pour les individus traités et si $k = 0$ on retrouve l'équation de résultat pour les individus non traités (groupe de contrôle).⁵⁷ X est une matrice de variables explicatives, ε_k le vecteur des termes d'erreurs suivant les hypothèses usuelles et β_k le vecteur des paramètres associés à X.

Les matrices W_k^1 et W_k^0 de dimension $(n \times n)$ permettent de spécifier les quatre formes de voisinages pouvant exister, à savoir, les interactions intra-groupes et inter-groupes et qui vont être combinées aux interactions spatiales spécifiées par W compte tenu de l'ordonnement des individus selon leur statut (B : bénéficiaire ou NB : non bénéficiaire). On peut visualiser la typologie des interactions, compte tenu d'un schéma spatial inhérent à W et à partir de laquelle sont déduites les matrices W_k^1 et W_k^0 (Figure 4-1) :

Figure 4-1. Matrice de voisinage et typologie des interactions

$$W = \begin{matrix} n_1 \\ n_0 \end{matrix} \left\{ \begin{array}{cc} \begin{array}{c} \text{Interactions intra-groupes} \\ \overbrace{B \times B} \\ \text{Interactions inter-groupes} \end{array} & \begin{array}{c} \text{Interactions inter-groupes} \\ \overbrace{B \times NB} \\ \text{Interactions intra-groupes} \end{array} \\ \hline \begin{array}{c} \overbrace{NB \times B} \\ \text{Interactions inter-groupes} \end{array} & \begin{array}{c} \overbrace{NB \times NB} \\ \text{Interactions intra-groupes} \end{array} \end{array} \right.$$

⁵⁷ L'estimation peut également être faite en introduisant dans l'équation de résultat global une distinction entre les individus bénéficiaires et les individus non bénéficiaires. Cette procédure reste difficile à mettre en œuvre, car la prise en compte des interactions inter et intra groupes impliquerait alors dans notre cas d'estimer quatre paramètres autorégressifs.

Au sein de l'équation de résultat pour les bénéficiaires ($k = 1$) on distinguera donc les interactions intra-groupes - entre les individus traités uniquement - avec W_1^1 et les interactions intergroupes - entre les individus traités et les non traités - avec W_1^0 . De la même manière, pour l'équation du groupe de contrôle ($k = 0$), W_0^0 identifie les interactions intra-groupes - entre les individus non traités uniquement - et W_0^1 les interactions inter-groupes - entre les individus non traités et les individus traités. Il s'agit donc ici d'une spécification qui permet de tenir compte de la diversité des voisinages, de leur caractère asymétrique et d'estimer les paramètres d'interaction ρ_{kl} associés ($k = 1$ ou 0 et $l = 1$ ou 0).

Les quatre matrices blocs sont représentées ci-après :

$$W_1^1 = \begin{pmatrix} B \times B & 0 \\ 0 & 0 \end{pmatrix} \quad W_1^0 = \begin{pmatrix} 0 & B \times NB \\ 0 & 0 \end{pmatrix} \quad W_0^1 = \begin{pmatrix} 0 & 0 \\ NB \times B & 0 \end{pmatrix} \quad W_0^0 = \begin{pmatrix} 0 & 0 \\ 0 & NB \times NB \end{pmatrix}$$

Si la matrice $(I_n - \rho_{k1}W_k^1 - \rho_{k0}W_k^0)^{-1}$ est inversible alors on peut écrire les résultats (cf. équation (4-14) sous la forme réduite suivante :

$$Y = (I_n - \rho_{k1}W_k^1 - \rho_{k0}W_k^0)^{-1}X\beta_k + (I_n - \rho_{k1}W_k^1 - \rho_{k0}W_k^0)^{-1}\varepsilon_k \quad (4-15)$$

d'espérance μ_k et de matrice de variance-covariance Ω_k :

$$\mu_k = (I_n - \rho_{1k}W_k^1 - \rho_{0k}W_k^0)^{-1}X\beta_k \quad (4-16)$$

$$\Omega_k = (I_n - \rho_{k1}W_k^1 - \rho_{k0}W_k^0)^{-1} \left[(I_n - \rho_{k1}W_k^1 - \rho_{k0}W_k^0)^{-1} \right]' \sigma_{\varepsilon_k}^2 \quad (4-17)$$

Les particularités de ces spécifications sont triples. D'une part, les effets des interactions endogènes (modèle SAR) sont modélisés pour explicitement prendre en compte les effets de diffusion de la politique sur les bénéficiaires attendus. Une spécification de type SEM, comme dans l'équation de sélection, aurait plutôt visé à tenir compte d'un problème d'autocorrélation spatiale pouvant être produit par tout un ensemble de raisons, mais sans fournir d'éléments pour comprendre le rôle des interactions sur les résultats. Une modélisation de type SLX (*Spatial Lag on Explanatory*), tenant compte des interactions exogènes sur les variables explicatives, permettrait effectivement de spécifier des effets de clubs selon les caractéristiques des individus, mais ne produit pas d'effets multiplicateurs (équation (4-18)).

La modélisation SAR semble plus utile à l'évaluation de la politique publique dans ce sens. D'autre part, les équations de résultats permettent d'étendre la prise en compte des interactions endogènes et de leurs effets aux situations inter-groupes là où Manski (2013), Hudgens et Halloran (2008) ou Rosenbaum (2007) ne considèrent que les interactions intra-groupes. Enfin, pour chaque individu, une double structure d'interactions est distinguée comme dans Lacombe (2004) pour estimer les effets *within* et *between* sur les résultats de la politique et ainsi sortir d'un cadre où les individus sont considérés comme isolés les uns des autres, ce que Lacombe (2004) spécifie comme des effets frontières. Lacombe (2004) explicite des effets de proximité d'un *county*, d'une part, avec les *counties* au sein du même Etat (situation *within*), et d'autre part, avec les *counties* les plus proches des Etats limitrophes (situation *between*). Dans notre cas, les effets *within* et *between* font référence aux interactions intra-groupes et inter-groupes et s'appliquent à des individus sélectionnés comme bénéficiaires ou non bénéficiaires.

L'estimation des paramètres spatiaux permet de statuer sur la présence d'effets multiplicateurs impactant les bénéficiaires ($\hat{\rho}_{10} \neq 0$ et/ou $\hat{\rho}_{11} \neq 0$) et les non bénéficiaires ($\hat{\rho}_{00} \neq 0$ et/ou $\hat{\rho}_{01} \neq 0$). Si les paramètres spatiaux sont de même signe, alors on peut, suivant la terminologie de Manski (2013), y associer des effets de renforcement (*reinforcing interactions*), ou des effets contraires (*opposing interactions*) si les paramètres sont de signes opposés.

Formellement, les effets de multiplicateur spatial sont estimés via la transformation spatiale $(I_n - \rho_{k1} W_k^1 - \rho_{k0} W_k^0)^{-1}$: toute variation d'une variable explicative X_p aura un effet différent de $\hat{\beta}_k$ sur la variable dépendante Y .⁵⁸

$$\frac{\partial Y_k}{\partial x_p} = (I_n - \rho_{k1} W_k^1 - \rho_{k0} W_k^0)^{-1} \beta_k \neq \beta_k \quad (4-18)$$

L'estimation de l'équation (4-15) par la méthode des moindres carrés produit des paramètres β_k biaisés et non convergents si les paramètres spatiaux ρ_{*k} sont significativement

⁵⁸ On parle également d'effet total ou d'effet de spillovers dont les écritures et la détermination, dans le cas d'un modèle SAR, sont présentées dans plusieurs travaux (voir par exemple, LeSage et Pace, 2009) et que nous ne détaillons pas dans le cadre de cet article.

différents de zéro. Pour estimer le modèle SAR avec deux matrices différentes, nous utilisons la méthode du maximum de vraisemblance (Lacombe, 2004).

La fonction de vraisemblance de l'équation (4-14) s'écrit comme suit : pour $k = 1 ; 0$:

$$L(\beta_k, \rho_{1k}, \rho_{0k}, Y_k, X, \sigma) = (2\pi\sigma^2)^{-n/2} |I_n - \rho_{k1}W_k^1 - \rho_{k0}W_k^0| \exp\left(-\frac{1}{2\sigma^2} \varepsilon_k' \varepsilon_k\right) \quad (4-19)$$

avec $\varepsilon_k = (I_n - \rho_{k1}W_k^1 - \rho_{k0}W_k^0)Y_k - X\beta_k = \mathbb{Y} - X\beta_k$

et $\mathbb{Y} = (I_n - \rho_{k1}W_k^1 - \rho_{k0}W_k^0)Y_k$

La fonction de log-vraisemblance est égale à :

$$\ln L = \ln |I_n - \rho_{k1}W_k^1 - \rho_{k0}W_k^0| - \frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{(\mathbb{Y} - X\beta_k)'(\mathbb{Y} - X\beta_k)}{2\sigma^2} \quad (4-20)$$

On obtient alors des estimateurs des paramètres β_k et σ_ε^2 :

$$\hat{\beta}_k = (X'X)^{-1}X'\mathbb{Y} \quad (4-21)$$

$$\hat{\sigma}_\varepsilon^2 = \frac{(\mathbb{Y} - X\hat{\beta}_k)'(\mathbb{Y} - X\hat{\beta}_k)}{n} \quad (4-22)$$

4.2.3. Identification de l'effet causal avec interactions endogènes

L'effet causal de la politique est mesuré par l'effet moyen du traitement sur les traités (Δ^{ATT} donné par l'équation (4-5)), ou en considérant ce que la politique aurait produit si elle avait été étendue à l'ensemble de la population, mesuré par l'effet global du traitement (Δ^{ATE} donné par l'équation (4-6)).

Cependant, l'effet du traitement sur les traités Δ^{ATT} n'est pas observable, car le contrefactuel $E(Y_0|T = 1, X, Z)$ est inobservable.

Ainsi, pour calculer l'effet moyen de traitement sur les traités, on recourt aux résultats observés des individus du groupe de contrôle, c'est-à-dire $E(Y_0|T = 0)$. On définit alors un nouveau paramètre Δ qui est l'estimateur de l'effet moyen du traitement sur les traités Δ^{ATT} .

$$\Delta = E(Y_1|T = 1) - E(Y_0|T = 0) \quad (4-23)$$

La différence entre l'effet moyen du traitement sur les traités Δ^{ATT} et son estimateur Δ représente le biais.

$$\begin{aligned} \Delta &= E(Y_1|T = 1) - E(Y_0|T = 1) + E(Y_0|T = 1) - E(Y_0|T = 0) \\ &= \Delta^{ATT} + E(Y_0|T = 1) - E(Y_0|T = 0) \\ &= \Delta^{ATT} + BS \end{aligned} \quad (4-24)$$

Le terme $[E(Y_0|T = 1) - E(Y_0|T = 0)]$ (équation (4-24)) représente le biais de sélection spatialement ajusté, qui est la différence entre le contrefactuel pour les personnes traitées et les résultats observés pour les personnes non traitées. Ce biais de sélection est présent lorsque la proxy, c'est-à-dire $E(Y_0|T = 0)$, n'est pas très proche du contrefactuel.

Si le biais de sélection est nul ($[E(Y_0|T = 1) - E(Y_0|T = 0)] = 0$) alors l'estimateur de l'effet moyen de traitement sur les traités Δ est égal à l'effet moyen du traitement sur la population qui est égal à la différence des résultats moyens observés pour les bénéficiaires et les non bénéficiaires.

$$\Delta = \Delta^{ATT} = \hat{\Delta}^{ATE} = E(Y_1|T = 1) - E(Y_0|T = 0) \quad (4-25)$$

Cependant, dans le cadre des politiques publiques non expérimentales où la sélection des bénéficiaires est non aléatoire, les paramètres d'intérêt (effet moyen de traitement sur les traités, par exemple) sont fortement exposés à la présence du biais de sélection, car il est évident que le résultat moyen du groupe de contrôle ne peut pas coïncider exactement avec le contrefactuel : $(E(Y_0|T = 1) - E(Y_0|T = 0)) \neq 0$. Dans ce type de politique où l'on prend en compte les interactions, il est nécessaire de contrôler le biais de sélection spatialement ajusté pour une bonne estimation de l'effet de la politique.

A partir des écritures de Schroeder (2010), les expressions du biais de sélection $E(\xi_1|T=1)$ et $E(\xi_0|T=0)$ sont données sous leur forme spatialement ajustée par les écritures suivantes :

$$E(Y_1|T=1) = \mu_1 + E(\xi_1|T=1) = \mu_1 + E(\xi_1|\nu > -Z\alpha) \quad (4-26)$$

$$E(Y_0|T=0) = \mu_0 + E(\xi_0|T=0) = \mu_0 + E(\xi_0|\nu < -Z\alpha) \quad (4-27)$$

avec : $\mu_1 = E(Y_1|T=1) = (I_n - \rho_{11}W_1^1 - \rho_{10}W_1^0)^{-1}X\beta_1$

et $\mu_0 = E(Y_0|T=0) = (I_n - \rho_{00}W_0^0 - \rho_{01}W_0^1)^{-1}X\beta_0$

(ξ_1, ν) et (ξ_0, ν) suivent une loi normale de moyenne 0 et de matrices de covariance respectives $\begin{pmatrix} \sigma_1 & \eta_1\sigma_1 \\ \eta_1\sigma_1 & 1 \end{pmatrix}$ et $\begin{pmatrix} \sigma_0 & \eta_0\sigma_0 \\ \eta_0\sigma_0 & 1 \end{pmatrix}$ où η_1 désigne le coefficient de corrélation, σ_1^2 la variance de ξ_1 , η_0 désigne le coefficient de corrélation, σ_0^2 est la variance de ξ_0 et, sans perte de généralité, le terme d'erreur ν est normalisé avec une variance égale à 1 (DiNardo et Lee, 2011).

Plus précisément :

$$\begin{aligned} E(\xi_1|T=1) &= E(\xi_1|\nu > -Z\alpha) = \eta_1\sigma_1 \frac{\phi(-Z\alpha)}{1 - \Phi(-Z\alpha)} \\ &= \eta_1\sigma_1 \frac{\phi(Z\alpha)}{\Phi(Z\alpha)} \end{aligned} \quad (4-28)$$

$$\begin{aligned} E(\xi_0|T=0) &= E(\xi_0|\nu < -Z\alpha) = \eta_0\sigma_0 \frac{-\phi(-Z\alpha)}{\Phi(-Z\alpha)} \\ &= \eta_0\sigma_0 \frac{-\phi(Z\alpha)}{1 - \Phi(Z\alpha)} \end{aligned} \quad (4-29)$$

Les équations (4-26) et (4-27) deviennent :

$$E(Y_1|T=1) = \mu_1 + \eta_1\sigma_1 \frac{\phi(Z\alpha)}{\Phi(Z\alpha)} \quad (4-30)$$

$$E(Y_0|T = 0) = \mu_0 + \eta_0\sigma_0 \frac{-\phi(Z\alpha)}{1 - \Phi(Z\alpha)} \quad (4-31)$$

En présence de biais de sélection dans un cadre spatial, l'effet causal de la politique devient finalement (équation (4-32)) :

$$\begin{aligned} \hat{\Delta}^{ATT} &= E(Y_1|T = 1) - E(Y_0|T = 0) = [\mu_1 + E(\xi_1|T = 1)] - [\mu_0 + E(\xi_0|T = 0)] \\ &= (\mu_1 - \mu_0) + \left[\eta_1\sigma_1 \frac{\phi(Z\alpha)}{\Phi(Z\alpha)} - \eta_0\sigma_0 \frac{-\phi(Z\alpha)}{1 - \Phi(Z\alpha)} \right] \\ &= (\mu_1 - \mu_0) + (\eta_1\sigma_1 - \eta_0\sigma_0) \frac{\phi(Z\alpha)}{\Phi(Z\alpha)} = \Delta^{ATE} + (\eta_1\sigma_1 - \eta_0\sigma_0) \frac{\phi(Z\alpha)}{\Phi(Z\alpha)} \end{aligned} \quad (4-32)$$

4.3. Application empirique

Pour mettre en œuvre le modèle et illustrer les implications des effets spatiaux sur l'évaluation des politiques publiques, nous réalisons une étude empirique sur données réelles. L'application choisie est la politique de revitalisation des zones rurales. Plusieurs éléments nous ont semblé intéressants dans cette politique, dont les principes de base sont rappelés dans un premier temps. En effet, elle concerne un ensemble assez important de communes rurales françaises dont la répartition géographique rend intéressante le traitement spatial des interactions entre communes. Cette politique est par ailleurs assez bien documentée, avec des données accessibles à l'échelle des communes ou à l'échelle de zones territoriales pertinentes pour l'analyse rurale comme les zones d'emplois. Un autre avantage est que la politique ne revêt pas un caractère obligatoire si bien qu'un groupe de contrôle peut être défini. Enfin, quelques études se sont intéressées à l'évaluation de cette politique (Lofredi, 2007) et notamment un travail récent mené par (Lorenceau, 2009). Ces deux études mobilisent des méthodes différentes d'évaluation de l'effet causal : une combinaison de la méthode d'appariement par le score de propension et de la méthode de double différence dans Lofredi (2007) et la méthode de discontinuité par la régression pour Lorenceau (2009). Ces approches, différentes de la nôtre et menées sur des périodes antérieures, concluaient à un effet non significatif de la politique sur le développement des cantons concernés. Cependant, aucune de ces études n'intègre les interactions dans l'estimation de l'effet causal et leurs résultats peuvent être questionnés si l'existence de ces interactions est considérée comme plausible et donc intégrée dans l'évaluation de la politique. Après avoir rappelé dans un premier temps, quelques éléments de cadrage sur le dispositif et le terrain d'étude, la méthodologie empirique est présentée dans une deuxième partie. Les résultats sont présentés dans la troisième partie et nous nous attacherons au-delà de l'illustration à donner quelques interprétations.

4.3.1. Cadrage de l'analyse : la politique des ZRR et échantillon d'étude

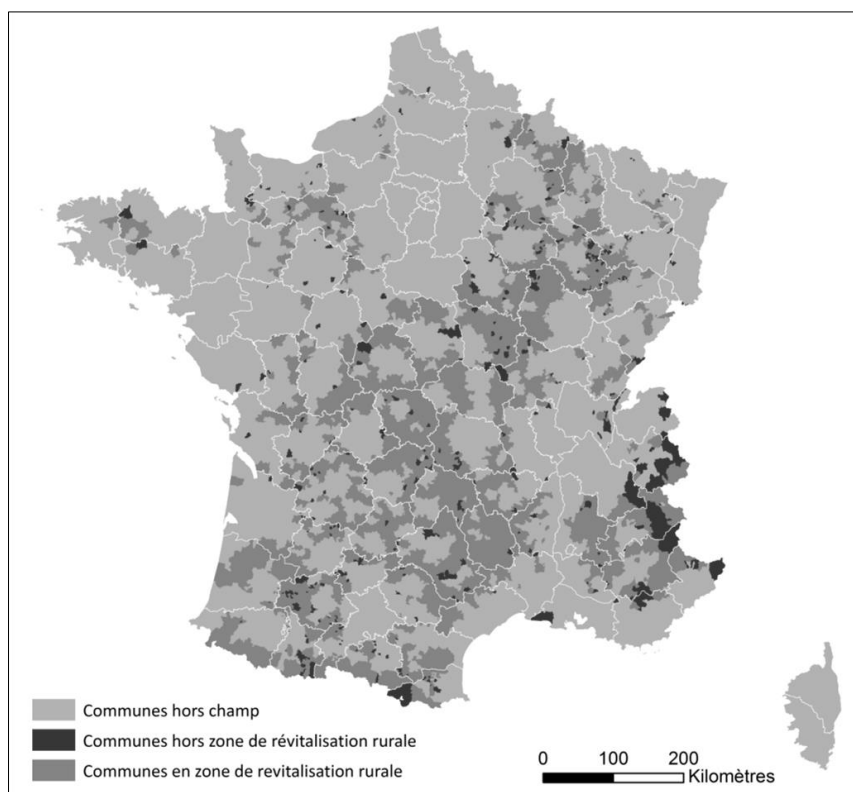
Les Zones de Revitalisation Rurale (ZRR), créées par la Loi d'Orientation pour l'Aménagement et le Développement du Territoire (LOADT) du 4 février 1995, désignent des territoires ruraux qui font face à des difficultés de développement. Cette loi vise à développer l'emploi et l'installation d'entreprises sur ces territoires. Elle permet aux acteurs économiques

- entreprises, artisans/commerçants, professions libérales, propriétaires immobiliers ... de ces zones de bénéficier d'un certain nombre d'avantages fiscaux tels que des exonérations d'impôt sur le revenu ou d'impôt sur les sociétés ou d'exonérations des taxes locales. L'éligibilité des communes à la politique de développement associée est fondée sur un principe territorial et sur un ensemble de trois critères : institutionnels, démographiques et socio-économiques dont la description précise est donnée en Annexe 1. Il s'agit de communes appartenant à des territoires (cantons, arrondissements ou EPCI) caractérisés par de faibles densités et présentant une forte représentation des activités agricoles ou une baisse de l'activité économique dont la liste est fixée par arrêté (cf. Annexe 1). En 2010, 13 060 communes ont été classées ZRR soit 36 % des communes de la France métropolitaine.⁵⁹

Pour constituer notre échantillon, nous avons considéré l'ensemble des communes de la France métropolitaine et avons exclu les communes de la région Corse afin de conserver une continuité physique du territoire. Parmi les communes restantes, nous avons retenu celles qui appartiennent à un canton dont la population est inférieure à 31 habitants/km². Ensuite, pour atténuer l'impact des grands pôles sur l'effet réel de la politique dans les communes rurales, nous avons également exclu les communes appartenant à l'espace des grandes aires urbaines qui est constitué des grands pôles urbains avec leurs couronnes ainsi que les communes multipolarisées des grandes aires urbaines (Insee, ZAU 2010). Au final, notre échantillon comporte 8 367 communes, dont 7 619 sont classées ZRR (au sens de l'arrêté du 31 décembre 2010) et 748 sont hors ZRR (ci-après dénommées HZRR). La répartition spatiale des communes permet de visualiser l'hétérogénéité des voisinages des communes selon leur statut (cf. Carte 4-1) : des communes classées ZRR sont entourées de communes ZRR mais également de communes HZRR comme dans les zones littorales ou de montagnes.

⁵⁹ Datar, Historique des communes classées ZRR, Arrêté du 10/07/2013. COG0910 pour le classement du 31 décembre 2010.

Carte 4-1 : Répartition spatiale des communes ZRR et hors ZRR de l'étude



Source : Datar, *Historique des communes classées ZRR, COG0910*. Cartographie MSH Dijon

Note : Sont exclues du champ d'analyse : les communes corses, les communes appartenant aux grandes aires urbaines (Insee, zonage en aire urbaine de 2010) et celles localisées dans un canton d'une densité supérieure à 31 habitants par km² en 2010.

Les communes ZRR sont les « bénéficiaires ». Les communes HZRR constituent le groupe de contrôle : bien que n'étant pas éligibles au dispositif, elles présentent des caractéristiques similaires à celles classées ZRR, en l'occurrence, elles appartiennent à un canton, dont la population est inférieure à 31 habitants/km².

Au total (cf. Tableau 4-1), les communes de notre échantillon représentent 29,5 % du territoire de la France métropolitaine et totalisent 2 817 042 habitants, soit 4,36 % de la population française (RGP 2010). Les communes ZRR comptent en moyenne 329 habitants contre 418 habitants en moyenne dans les communes HZRR. Le taux de chômage moyen dans les communes ZRR est de 9,84 % en moyenne et il est légèrement supérieur au taux de chômage au niveau national (9,4% en 2010). Il est en revanche plus élevé que celui des communes HZRR (8,1 % en moyenne). Par ailleurs, l'emploi agricole représente en moyenne 12,8 % des emplois dans les communes ZRR contre 8 % dans les communes hors ZRR.

Le développement économique des communes rurales est apprécié dans notre analyse par la création d'établissements en 2011 (cf. Tableau 4-1) : 25 230 établissements ont été créés dans les communes considérées et 84 % des créations ont été faites dans les communes ZRR. Les créations ont été les plus soutenues dans les secteurs du commerce, des transports et des services divers (63%) aussi bien dans les communes ZRR que les communes HZRR. On note en revanche que dans le secteur « administration, santé et enseignement », la création d'établissements est de 16 % pour les communes HZRR contre 9 % dans les communes ZRR. A l'inverse, pour les communes HZRR, 7 % des établissements sont créés dans le domaine industriel contre 13 %, soit près du double, pour les communes ZRR.

Tableau 4-1 : Caractéristiques des communes selon leur statut

Statut des communes (en 2010)		HZRR	ZRR	Total
Population communale (2010)	Moyenne	418	329	337
	Total	312 689	2 504 353	2 817 042
Taux de chômage moyen (2010)		8,1%	9,8%	9,6%
Emploi agricole (2010)		8%	12,78%	
Création d'établissements (2011)				
Total		4 008	21 222	25 230
Par secteurs d'activités				
<i>Industrie</i>		299	2 755	3 054
<i>Construction</i>		528	3 148	3 676
<i>Commerce, Transport, Services</i>		2 537	13 352	15 889
<i>Administration, santé, enseignement</i>		644	1 967	2 611

Source : Insee, RGP (2010) et Chiffres clés - Démographie des entreprises 2006 2011, Sirene. Calcul des auteurs.

4.3.2. Méthodologie de l'analyse empirique

L'objectif de l'application est d'estimer l'effet du classement de la commune en ZRR sur la création d'établissements, en contrôlant l'indépendance entre le classement et le résultat (biais de sélection) et en prenant en compte les interactions entre les communes, quel que soit leur statut : ZRR ou HZRR. Nous présentons les variables dépendantes et exogènes et les matrices d'interactions utilisées dans l'application empirique. L'ensemble des précisions sur les bases de données, les mesures et les questions économétriques sont détaillées en Annexe 2.

Les variables dépendantes. Pour l'équation de sélection, la variable endogène ZRR est une variable binaire, qui est égale à 1 si la commune est classée ZRR (en 2010) et 0 sinon.

Pour les équations de résultats, nous considérons le nombre d'établissements créés en 2011. La variable est exprimée en logarithme.

Les variables exogènes. Le Tableau 4-2 fournit une présentation synthétique des variables utilisées dans l'équation de sélection, d'une part, et dans les équations de résultats, d'autre part.

Le processus de sélection va dépendre d'un ensemble de caractéristiques Z pour partie liées aux critères d'éligibilité. Au titre des critères de classement, nous considérons la densité du canton (Dpop), la population municipale (PopCom), la population active au chômage (CHOM) et la population des agriculteurs (AGRI). Nous considérons par ailleurs une variable permettant de catégoriser les communes en fonction du zonage en aire urbaine de 2010 et ceci afin de capter l'effet d'une polarisation éventuelle sur les communes de l'échantillon. La variable AIRU comprend plus précisément 4 modalités (cf. Tableau 4-3) permettant de repérer l'appartenance aux « moyennes aires urbaines », aux « petites aires urbaines »⁶⁰, ou le fait d'être dans le groupe des « autres communes multi-polarisées » ou des « communes isolées, hors influence des pôles ». La répartition des communes de l'échantillon nous montre que les communes ZRR appartenant à des aires urbaines sont surtout présentes dans les groupes de communes hors influence d'une aire urbaine (en tant que commune multipolarisée ou commune isolée) mais que 7 % d'entre elles sont situées dans les aires urbaines. Cette représentation est quasiment identique pour les communes de contrôle HZRR : 8 % d'entre elles seulement sont sous l'influence d'une aire urbaine.

Pour les équations de résultats, nous avons considéré les variables explicatives de création d'établissements associées à différents mécanismes d'attractivité ou de performance (cf. Le Tableau 4-2), toutes proportions gardées du fait que nous nous situons dans des espaces ruraux et non urbains.

Ainsi, la population de la commune permet de tenir compte d'un effet de « taille de marché » notamment pour les activités de services ou de commerces. De même, le type de la commune selon le zonage en aire urbaine décrit précédemment est retenu. La structure de l'emploi, en termes de catégories socio-professionnelles, permet d'apprécier un effet

⁶⁰ Les petites ou moyennes aires urbaines se distinguent selon la taille du pôle urbain : entre 1 500 et 5 000 emplois pour les petites aires urbaines et entre 5 000 et 10 000 emplois pour les moyennes aires urbaines (Insee Méthodes, 2011).

d'appariement de la main d'œuvre sur les nouveaux établissements. Cette structure est mesurée au niveau du bassin d'emplois, à la fois pour des raisons statistiques et d'appréciation du marché de l'emploi. La population active municipale au chômage complète cette description du marché de l'emploi. Pour tenir compte d'effets potentiels de structures économiques liés au tissu d'activités économiques, nous avons pris en compte les catégories d'activités économiques à l'échelle du bassin d'emplois.⁶¹

⁶¹Cette échelle constitue actuellement le niveau d'analyse le plus pertinent pour les problématiques des marchés d'emploi locaux mais aussi pour des questions de validité et fiabilité des résultats, car dans certaines communes ces emplois sont très faibles voire inexistantes.

Tableau 4-2 : Variables exogènes du modèle d'évaluation⁶²

Variables explicatives dans l'équation de sélection	
Abréviation	Intitulé (unité)
DPop	Densité au niveau canton, 2010 (hab./km ²)
PopCom	Population de la commune en 2010 (en log)
AGRI	Population active agricole de la commune en 2010 (en log)
CHOM	Population active au chômage de la commune en 2010 (en log)
MultiPol	Autres communes multi-polarisées (dummy, <i>référence</i>)
Isolée	Communes isolées, hors influence des pôles (dummy)
AUMoy	Moyennes aires urbaines (dummy)
AUPetit	Petites aires urbaines (dummy)
Variables explicatives dans les équations de résultats	
Abréviation	Intitulé (unité)
PopCom	Population de la commune en 2010 (en log)
CHOM	Population active au chômage de la commune en 2010 (en log)
AV-2005	Communes classées en ZRR avant 2005 (= 0 à partir de 2005)
MultiPol	Communes multi-polarisées (dummy, <i>référence</i>)
Isolée	Communes isolées, hors influence des pôles (dummy)
AUMoy	Moyennes aires urbaines (dummy)
AUPetit	Petites aires urbaines (dummy)
ARTISANS*	Population active de la CSP Artisans (zone d'emploi, 2010, racine carrée)
EMPLOYES*	Population active de la CSP Employés (zone d'emploi, 2010, racine carrée)
EMP-AGRI*	Emplois au lieu de travail dans l'agriculture (zone d'emploi, 2010, racine carrée)
EMP-INDUS*	Emplois au lieu de travail dans l'industrie (zone d'emploi, 2010, racine carrée)
EMP-ADMIN*	Emplois au lieu de travail dans l'administration (zone d'emploi, 2010, racine carrée)
TYPOLOGIE DES PAYSAGES DES CAMPAGNES	
10 modalités (cf. Tableau 4-4) : variable dummy	
Modalité de référence : Plaines et plateaux, prairies et forêts (PP_PF)	

* Les variables retenues sur les catégories d'emplois et d'activités ont été sélectionnées par une procédure *stepwise*

Afin de contrôler un éventuel effet d'ancienneté et de lisibilité du dispositif ZRR sur la création des établissements, nous avons distingué les communes qui sont entrées dans le dispositif avant 2005 ou après 2005 ; année qui correspond à un complément institutionnel du dispositif (appartenance à un EPCI notamment) et à partir de laquelle des listes régulières des

⁶² Les variables ARTISAN, EMPLOYES, EMP-AGRI, EMP-INDUST, EMP-ADMIN ont été transformées par la racine carrée pour les normaliser, car elles étaient fortement asymétriques à droite.

communes classées sont disponibles.⁶³ 854 communes de notre échantillon sont ainsi entrées dans le dispositif ZRR, ce qui représente 11 % des communes ZRR de l'échantillon. Enfin, nous avons tenu compte d'un effet possible des aménités naturelles sur l'attractivité des nouveaux établissements et avons utilisé la typologie *des campagnes françaises et des espaces à enjeux spécifiques* (Hilal et al., 2011) qui comprend 10 modalités pour décrire si la commune est plus ou moins artificialisée, située sur le littoral ou en montagne, située en plaine cultivée, marquée par des reliefs, composée de paysages forestiers.... Le Tableau 4-3 précise la répartition des communes selon chacune des modalités. Une procédure de sélection de type *stepwise* a permis de retenir les variables d'emplois et d'activités les plus significatives, à savoir les artisans et les employés et les activités de commerces, industrielles et administratives.

Tableau 4-3 : Typologies thématiques et répartition des communes

Typologie des aires urbaines*	Libellé	Hors ZZR	ZRR	Total
Communes isolées, hors influence des pôles	Isolée	429	4852	5281
Communes multi polarisées	MultiPol	258	2231	2489
Moyennes Aires urbaines	AUMoy	27	269	296
Petites Aires urbaines	AUPeti	34	267	301
Total		748	7619	8367

Typologie des paysages**	Libellé	Hors ZZR	ZRR	Total
Rural artificialisé, présence forte du bâti	RA_FB	12	81	93
Paysages artificialisés, littoraux	PA_litt	4	13	17
Paysages de cultures, forte marque du bâti	PC_FB	24	341	365
Plaines et plateaux et grands openfields	PP_GOF	41	352	393
Plaines et plateaux, campagne diversifiée avec bâti marqué	PP_CDB	89	1016	1105
Plaines et plateaux, prairies et forêts	PP_PF	36	399	435
Collines et vallées, paysages diversifiés	CV_PD	78	986	1064
Collines et vallées en prairies, peu artificialisées	CVP_PA	144	2060	2204
Relief marqué, paysages diversifiés	RM_PD	91	1221	1312
Montagne, très forts reliefs et végétation semi-naturelle	MR_VSN	229	1150	1379
Total		748	7619	8367

* Source : Insee, *Nouveau Zonage en Aires Urbaines 2010*, calculs des auteurs

** Source : Datar, *Typologie thématique du champ « Paysages », 2011*, calculs des auteurs.

⁶³ Avant 2005, il n'est en fait pas possible de calculer une ancienneté réelle des communes dans le dispositif ZRR. Les autres études évaluant la politique ZRR (Lofredi, 2007 ou Lorenceau, 2009) n'intègrent pas non plus cette information.

Les matrices de voisinage

La structure des interactions entre les communes est modélisée par les matrices de voisinage de telle sorte qu'une commune *i* interagit avec toutes les communes *j* situées dans un rayon de 100 kilomètres. Cette distance est couverte facilement de manière quotidienne et toutes les communes sont en interaction avec au moins une autre commune. Les valeurs des termes w_{ij} sont données par l'inverse de la distance entre les communes voisines au sens de ce seuil et sont égales à 0 pour les communes éloignées de plus de 100 kilomètres. La matrice *W* (8367×8367) est créée de telle sorte que les communes ZRR occupent les 7 619 premiers rangs et les communes HZZR les 748 rangs suivants. Dans ces conditions, la partition de la matrice *W* en quatre quadrants permet de repérer les quatre types d'interactions (cf. Figure 4-2).

Figure 4-2 : Matrice de voisinage et typologie des interactions

$$W = \begin{pmatrix} \text{Interactions intra-groupes} & \text{Interactions inter-groupes} \\ \overbrace{ZRR \times ZRR} & \overbrace{ZRR \times HZZR} \\ \hline \overbrace{HZZR \times ZRR} & \overbrace{HZZR \times HZZR} \\ \text{Interactions inter-groupes} & \text{Interactions intra-groupes} \end{pmatrix}$$

Le tableau 4-4 indique quelques caractéristiques d'interactions, selon que l'on considère la totalité des communes ou seulement les quadrants concernés par les voisinages intra-groupes ou inter-groupes. Nous remarquons que les interactions intra-groupes sont en moyenne importantes, 608 pour les communes ZRR et 482 pour les communes hors ZRR, mais pas radicalement différentes, selon le type de commune, en dépit du nombre très différent de communes dans chaque catégorie. Le choix du seuil de distance, support des interactions, permet de ce fait de préserver l'organisation spatiale réelle des communes. Nous remarquons par ailleurs que les interactions inter-groupes restent importantes, une commune d'un groupe interagissant en moyenne avec une cinquantaine de communes d'un autre groupe.

Tableau 4-4 : Caractéristiques des matrices de voisinage

	Min	Max	Moyenne
W	21	950	479,16
<i>Interactions intra-groupes</i>			
Communes ZRR avec des voisins ZRR (ZRR-ZRR)	30	1 122	608
Communes HZZR avec des voisins HZZR (HZZR-HZZR)	37	1 045	482
<i>Interactions inter-groupes</i>			
Communes ZRR avec des voisins HZZR (ZRR-HZZR)	1	120	47,32
Communes HZZR avec des voisins ZRR (HZZR-ZRR)	5	124	57,67

4.3.3. Résultats des estimations⁶⁴

Pour chaque étape du modèle d'évaluation, nous comparons les résultats obtenus par notre modèle spatial avec ceux obtenus par une modélisation dite standard ou a-spatiale, c'est-à-dire sans tenir compte des interdépendances. A travers ces estimations, nous donnons quelques interprétations des effets du programme ZRR sur le développement des communes.⁶⁵

Equation de sélection.

Le Tableau 4-5 présente les résultats de l'estimation de l'équation de sélection avec un modèle Probit standard (colonne 1) et le modèle Probit spatial (colonne 2).⁶⁶

L'estimation standard de l'équation de sélection montre que les critères d'éligibilité en ZRR sont déterminants et avec les signes attendus : les communes appartenant à des cantons avec de fortes densités et les communes avec une population municipale importante ont moins de chance d'être en ZRR. En revanche, les caractéristiques de la population active sont significatives et positives. En d'autres termes, les communes avec de fortes proportions de population agricole et de population au chômage ont plus de chance d'être en ZRR. Enfin, par rapport au zonage en aires urbaines, les estimations montrent que les communes appartenant aux communes isolées, hors influence des pôles, sont plus susceptibles d'être en ZRR par rapport aux autres communes multipolarisées. Ces résultats sont sensiblement modifiés lorsque l'on tient compte des interdépendances entre les communes, avec trois implications principales. Premièrement, le paramètre spatial λ est positif et significatif (0,698) : les interdépendances entre les communes existent du fait de caractéristiques inobservables. La structure des voisinages n'est pas neutre quant à la possibilité qu'une commune soit ou non en ZRR et le paramètre spatial peut traduire l'effet de la diffusion d'un choc aléatoire au sein des communes sur leur décision de participer. Deuxièmement, les caractéristiques des communes en termes de densités et de chômage restent significatives et du signe attendu alors que les

⁶⁴ Les estimations standards ont été faites sous R avec le package *SampleSelection*. Pour les estimations spatiales, l'équation de sélection est estimée sous Matlab (Toolbox J. Lesage). Les estimations des équations de résultats ont été faites sous Matlab (routines D. Lacombe utilisées dans l'article LACOMBE, 2004). Les calculs ont été effectués en utilisant les ressources du CHP du DSI-CCUB (Université de Bourgogne).

⁶⁵ Dans la mesure où une seule matrice W est employée (distance 100 km), les résultats restent contingents à cette structure de voisinage. Nous rappelons qu'une bonne pratique méthodologique, consisterait à réaliser une analyse de la qualité des estimations et de la sensibilité des résultats à plusieurs matrices de poids (variation de la distance, variation des critères de contiguïté...). Ces analyses sont en dehors de l'objectif de notre illustration empirique.

⁶⁶ Les estimations des effets marginaux sont disponibles auprès des auteurs.

autres variables (PopCom, AGRI et Isolée) ne sont plus déterminantes dans le processus de sélection. La présence d'autocorrélation spatiale des erreurs rend inefficients les estimateurs standards des paramètres. Au seuil $p = 0,8$, le modèle standard (non spatial) a classé comme ZRR 99 % des communes qui le sont réellement et comme HZRR, 97 % des communes qui le sont réellement. Ces pourcentages sont de 100 % pour le modèle spatial.

Tableau 4-5 : Résultat de l'estimation de l'équation de sélection

Variable dépendante : classement ZRR ou HZRR

Nombre d'observations : 8 367

	Probit standard coefficient [1]	Probit spatial coefficient [2]
CONSTANTE	2,484***	0,965***
DPop	-0,015***	-0,003***
PopCom	-0,294***	-0,001
CHOM	0,206***	0,009.
AGRI	0,100***	0,000
AUMoy	0,108	-0,017
Isolée	0,078.	0,007
AUPetit	0,023	0,005
MultiPol	<i>ref</i>	<i>Ref</i>
Lambda		0,698***
% de bonnes prédictions « ZRR » (au seuil de 0,8)	99%	100%
% de bonnes prédictions « HZRR » (au seuil de 0,8)	97%	100%
% de bonnes prédictions	99%	100%

Significativité : *** : 0,001 / ** : 0,01 / * : 0,05 / . : 0,1

L'estimation des paramètres des équations de sélection permet de calculer l'Inverse du Ratio de Mills standard (IRM) et l'Inverse du Ratio de Mills spatialement ajusté (IRMS) pour contrôler le biais de sélection dans les équations de résultats.

Equations de résultats

Le Tableau 4-6 présente respectivement les estimations des équations de résultats dans les communes ZRR et hors ZRR sans tenir compte des interactions (modèle a-spatial estimé par les MCO) ou avec les interactions (modèle SAR estimé par ML, Lacombe, 2004). Le test de dépendance spatiale globale effectué sur les créations d'établissements avec la matrice W indique une autocorrélation spatiale positive et significative : la statistique du I de Moran est égale à 0,121 (écart-type = 0,00118).

Tableau 4-6 : Estimations des résultats pour les communes ZRR et hors ZRR

Variable dépendante : création d'établissements (en log)

Nombre d'observations : 8 367

Variables	Communes ZRR		Communes HZRR	
	Modèle a-spatial	Modèle SAR	Modèle a-spatial	Modèle SAR
	MCO [1]	ML [2]	MCO [3]	ML [4]
CONSTANTE	-1,579***	-0,991***	-3,853***	-0,671**
PopCom	0,463***	0,390***	0,738***	0,055***
CHOM	0,108***	0,123***	-0,009	0,005
ARTISAN	0,017***	-0,006**	0,013	-0,003**
EMPLOYES	-0,024***	-0,006.	0,023**	0,007***
EMP-AGRI	-0,002	-0,001	-0,006.	0,000
EMP-INDUST	0,001	0,001	-0,011***	-0,001*
EMP-ADMIN	0,015***	0,008***	-0,016*	-0,004***
An-2005	0,471***	0,161***		
AUMoy	0,035	0,052	-0,317**	-0,010
Isolée	0,068***	0,037**	0,045	-0,005
AUPetit	0,054	0,043	0,018	0,002
MultiPol	Ref	Ref	ref	Ref
<i>Variables d'aménités : typologie des paysages⁽¹⁾</i>				
RA_FB	0,261***	0,376***	0,268	0,059*
PA_litt	0,224	0,286**	0,364	0,099
PC_FB	-0,041	0,033	0,097	0,032.
PP_GOF	0,021	0,067.	0,041	0,039*
PP_CDB	0,021	0,119***	-0,043	0,027.
PP_PF	ref	Ref	ref	Ref
CV_PD	0,002	0,110***	0,068	0,029*
CVP_PA	0,007	0,069**	0,135	0,048**
RM_PD	0,082*	0,127***	0,087	0,035*
MR_VSN	0,144***	0,194***	0,269*	0,052***
IRM ou IRMS	-2,515***	-4,404***	0,412*	0,242
Interactions intra- groupes	ZRR ρ_{11}	0,731***	HZRR ρ_{00}	0,957***
Interactions inter- groupes	ZRR-HZRR ρ_{10}	0,397***	HZRR-ZRR ρ_{01}	0,284***
Test d'égalité des coefficients (Wald)		0,117		3,412
R ² ou Pseudo-R ²	0,498	0,556	0,713	0,703
R ² Adj. ou Pseudo- R ² Adj.	0,497	0,552	0,706	0,703

(1) Pour la signification des variables d'aménités, se reporter au Tableau 4-3

Significativité : *** : 0,001 / ** : 0,01 / * : 0,05 / . : 0,1

Les créations d'établissements suivent une tendance spatiale similaire au sein des communes et ceci sans a priori sur le statut de ces communes en tant que bénéficiaires ou non de la politique ZRR. Dans la mesure où la spécification SAR considère les interdépendances endogènes entre les individus, nous pouvons interpréter ces interactions. Nous analysons plus particulièrement l'impact qualitatif de la prise en compte des interdépendances sur les résultats en comparant les estimations standards et spatiales des deux modèles.⁶⁷

Nous analysons tout d'abord les résultats pour les communes bénéficiaires (cf. Tableau 4-6, colonnes 1 et 2). Pour les variables explicatives, nous remarquons que les créations d'établissements sont, dans les deux cas, significativement et positivement influencées par la taille de la commune (PopCom) et son niveau de chômage (CHOM). L'ancienneté des communes dans le dispositif ZRR a un impact significativement positif sur la création d'établissements. De même, les communes isolées en dehors de l'influence d'une zone urbaine sont plus favorables à la création d'établissements que les communes multipolarisées sous influence de plusieurs zones urbaines. Plusieurs types d'aménités paysagères influencent positivement la création d'établissements dans les communes ZRR avec une tendance assez forte pour les paysages plutôt artificialisés (RA_FB), du littoral (PA_litt), de la campagne (CVP_PA) ou marqués par les reliefs (RM_PD, MR_VSN). Une structure administrative du bassin d'emploi impacte positivement la création d'établissements, mais de manière deux fois moindre une fois les interdépendances considérées, traduisant la complémentarité territoriale entre les communes à cette échelle. Ne pas considérer les interdépendances conduirait à surestimer le tissu économique de la commune. De même, des effets de concurrence sont plutôt repérés sur l'activité artisanale une fois les interdépendances prises en compte : plus cette catégorie est représentée dans le bassin d'emploi et moins la création d'établissements sera forte. Ne pas prendre en compte les interactions entre communes conduirait à considérer une influence positive de l'activité artisanale à l'échelle de chaque commune. Les estimations des paramètres d'interactions ρ_{11} entre les communes classées ZRR et ρ_{10} entre les communes classées ZRR et leurs voisins hors ZRR permettent de compléter ces analyses. Les valeurs estimées sont positives et significatives dans les deux situations : la création d'établissements dans une commune est positivement influencée, au-delà de ses propres

⁶⁷ Dans la mesure où les effets des variables exogènes du modèle SAR ne sont pas directement interprétables par la valeur estimée des coefficients, nous ne les comparerons pas avec les valeurs estimées par les MCO du modèle a-spatial. Les effets totaux du modèle SAR ne sont pas présentés, car leur calcul n'est pas nécessaire à l'évaluation de l'effet causal, qui est l'objectif du modèle d'évaluation de la politique dans cette illustration.

caractéristiques, par la création moyenne d'établissements dans les communes voisines, que celles-ci soient bénéficiaires ou non de la politique ZRR. Les interdépendances intra-groupes, entre communes ZRR, sont plus fortes que les interactions intergroupes. Il existe bien un impact différencié selon le type de voisinage, les valeurs des paramètres étant significativement différentes. Dans la mesure où les deux coefficients d'interactions sont de même signe, on est, cependant, dans le cas d'un renforcement des impacts.

Ces résultats montrent finalement que la politique ZRR semble être en phase avec ses objectifs de relance des activités économiques par le biais de la création d'établissements dans les communes rurales. Les aménités paysagères jouent un rôle attendu notamment pour les communes bénéficiant d'aménités touristiques. Notre modélisation permet, cependant, de montrer la présence d'interdépendances entre les communes rurales dont les effets sur les estimations sont de trois types. Premièrement, les estimateurs MCO sont biaisés et non convergents ce qui conduit à des résultats incorrects sur l'influence de certains déterminants. Deuxièmement, des effets différenciés selon les types d'interdépendances sont mesurés qui reflètent l'hétérogénéité des voisinages. Troisièmement, le paramètre associé au biais de sélection IRMS dans le cas spatial est significatif et indique qu'il faut tenir compte du biais de sélection dans l'estimation des effets de traitement. Nous rappelons que ce biais de sélection, comprend dans notre cas, les effets des interdépendances sur les caractéristiques inobservables du processus de classement des communes.

Pour l'estimation des créations d'établissements dans les communes hors ZRR, les résultats sont donnés dans les colonnes 3 et 4 du Tableau 4-6 lorsque les interdépendances entre communes sont négligées (modèle a-spatial) ou non (modèle SAR). Les paramètres spatiaux, d'interactions ρ_{00} entre les communes classées HZRR et ρ_{01} entre les communes HZRR et leurs voisines classées ZRR sont significatifs et positifs, avec un effet très fort pour les interactions intra-groupes. Il existe bien un impact différencié selon le type de voisinage, les valeurs des paramètres étant significativement différentes. Pour autant, les créations d'établissements dans une commune et dans les communes voisines sont positivement corrélées, avec les deux types de voisinages. On est dans le cas d'effet de renforcement des interactions. La prise en compte des interdépendances entre les communes semble de surcroît importante, car l'effet significatif du biais de sélection dans le modèle a-spatial n'est pas confirmé lorsque les interdépendances spatiales sur les inobservables sont prises en compte :

le coefficient estimé IRMS n'est pas significatif. L'estimation du modèle SAR montre que l'effet taille des communes ainsi que la présence d'employés dans la zone d'emplois sont des déterminants positifs de la création d'établissements. Les aménités paysagères jouent favorablement pour les paysages avec des reliefs marqués, ce qui est la marque assez distinctive dans notre échantillon des communes HZRR situées dans les territoires de montagne (cf. Figure 4-2). Lorsque l'on tient compte des interdépendances, la composition économique des bassins d'emplois joue négativement ce qui peut traduire un effet de concurrence territoriale associé aux interactions.

Si nous comparons les processus pour les communes ZRR et HZRR, nous soulignons des effets communs et différents. Les paramètres spatiaux ρ_{kl} sont significativement positifs ce qui souligne l'hétérogénéité des voisinages avec la présence d'interactions inter et intra-groupes. Suivant la littérature, on peut également parler d'effets de renforcement (Manski, 2013) entre les communes sous la forme de spillovers spatiaux ou de multiplicateur spatial (LeSage et Pace, 2009). Cet effet d'entraînement est par ailleurs a-symétrique puisque les interactions inter-groupes ne sont pas d'ampleur équivalente : les voisinages composés de communes HZRR impactent plus fortement les communes ZRR que ne le font les voisinages composés de communes ZRR pour les communes HZRR. On remarque néanmoins que la création d'établissements dans les communes hors ZRR est plus poussée par la taille de la commune et son économie. L'influence du zonage urbain n'est pas significative pour les communes hors ZRR. En revanche, pour les communes classées en ZRR, les déterminants liés à la revitalisation de l'économie hors influence des aires urbaines semblent pousser la création d'établissements.

Ces différents résultats peuvent maintenant être mis en correspondance par l'estimation de l'effet causal de la politique.

Estimation des effets de traitement

L'effet causal du classement ZRR sur la création d'établissements, avec ou sans interactions entre les communes, est estimé par l'effet moyen du traitement sur les traités Δ^{ATT} et par l'effet moyen du traitement sur la population Δ^{ATE} . Les résultats sont présentés dans le cas standard et dans le cas spatial (cf. Tableau 4-7). Le cas spatial correspond à une situation où l'endogénéité du traitement considère l'impact potentiel des interactions, ce qui correspond à un élargissement du domaine d'invalidité de l'hypothèse CIA, et à une situation

de relâchement de l'hypothèse SUTVA. Le cas standard correspond à une situation où ni l'endogénéité du traitement, ni les objectifs assignés à la politique ne prennent en compte les effets potentiels des interactions entre communes. Il est alors implicitement supposé dans ce cas, une validité partielle de l'hypothèse CIA et la validité de l'hypothèse SUTVA.

Tableau 4-7 : Estimations des effets de traitements

		Modèle a-spatial		Modèle spatial	
		ATT	ATE	ATT	ATE
nb	d'établissements	2,675..	2,445..	3,7	3,6
estimés					
St.dev		1,503	1,427	4,1	4,05
Biais (%)		8,7%		1,2%	

Significativité : .. : 0,1

Si on ne prend pas en compte les interactions entre les communes, alors on conclut que cette politique a effectivement un effet positif et significatif (au seuil de 10%). La création d'établissements dans les communes traitées peut être considérée comme supérieure à celle que l'on aurait obtenue si ces communes n'avaient pas bénéficié des actions liées à la politique (+2,7). De même, la création d'établissements du fait des dispositifs liés à la politique des ZRR peut être considérée comme supérieure pour l'ensemble des communes à ce qu'elle aurait été sans cette politique (+2,4). Cependant, la significativité des paramètres d'intérêt est à prendre avec précaution dans la mesure où nous avons montré dans les estimations des équations de sélection et de résultats qu'il y a respectivement de l'autocorrélation des erreurs et de l'autocorrélation dans les variables dépendantes. Dès lors, *toutes choses égales par ailleurs*, les paramètres d'intérêts ATT et ATE obtenus à partir de ces estimateurs sont biaisés. Par ailleurs, dans un cadre non spatial, la différence entre les paramètres ATT et ATE indique une distorsion de l'ordre de 9 % assignée au biais de sélection non contrôlé.

Ainsi, les estimations des effets moyens du traitement lorsque les interdépendances sont prises en compte, montrent un tout autre résultat. Il n'est pas possible dans ce cas d'imputer aux dispositifs de la politique ZRR un effet sur la création d'établissements dans les communes classées ou au sein de l'ensemble des communes, car les estimations des effets ATE et ATT sont positives mais ne sont plus significatives. Ainsi, même si les créations d'établissements dans ces communes sont avérées, la politique ne peut être considérée comme

une cause différenciant ce qui se passe entre une commune ZRR ou une autre commune similaire. Le modèle spatial est plus performant quant au contrôle du biais de sélection avec seulement 1,2 % de distorsion sur les paramètres ATT et ATE.⁶⁸

Nos conclusions penchent plus vers un effet mitigé du classement des communes en ZRR. Nous rejoignons ainsi les résultats obtenus par Lorenceau (2009), mais en restant prudent quant à la confirmation de ses résultats par les nôtres du fait d'une échelle d'analyse différente (canton et non commune) et de données différentes (2007 contre 2010).

⁶⁸ L'importance du contrôle du domaine de validité des hypothèses CIA ou SUTVA peut être appréciée en estimant par exemple l'effet causal lorsque l'on contrôle les interactions uniquement sur l'endogénéité du traitement ou uniquement sur les résultats attendus de la politique. Nous avons conduit les estimations dans ces deux situations et les résultats indiquent que l'effet causal est plutôt biaisé par l'absence de prise en compte des interactions dans l'estimation des résultats que dans le processus de sélection (les résultats sont disponibles auprès des auteurs).

Conclusion

Dans un cadre non expérimental, un ensemble d'hypothèses sont généralement postulées afin de contrôler la portée des interdépendances entre les individus sur les effets de la politique publique. Dans cette recherche, nous avons proposé de revenir sur la validité des hypothèses CIA (*Conditional Independence Assumption*) et SUTVA (*Stable Unit Treatment Value Assumption*) que l'on trouve en particulier combinées dans les modèles d'évaluation des effets de traitement en deux étapes. En mobilisant les modèles économétriques spatiaux, nous avons souligné trois domaines de réflexion. Premièrement, nous avons considéré que les décisions de participation des individus aux politiques pouvaient être impactées par des interactions dites corrélées au sens de Manski (1993) et nous avons proposé une spécification à erreurs spatialement autocorrélées du modèle probit de sélection. Deuxièmement, les interactions entre les individus peuvent se situer dans des voisinages hétérogènes au regard de leur statut de bénéficiaires ou de non bénéficiaires de la politique. Nous avons ainsi étendu le périmètre des hypothèses SUTVA et PIA, plutôt confiné aux interactions intra-groupes, aux cas des interactions inter-groupes. Nous pouvons dans ce cadre différencier les interactions entre les individus et en estimer la portée dans les résultats obtenus. Troisièmement, les équations de résultats ont été modélisées sous la forme de processus spatialement autorégressifs intégrant l'hétérogénéité des voisinages et une correction du biais de sélection par l'Inverse du Ratio de Mills, spatialement ajusté.

La politique territoriale de développement des zones rurales a été choisie pour illustrer empiriquement les extensions contenues dans notre modèle. L'évaluation des effets du classement des communes en Zone de Revitalisation Rurale a été effectuée sur les créations d'établissements. Les résultats confirment la présence d'effets de voisinage différenciés sur les créations d'établissements selon que les communes et leurs voisines sont classées ou non en ZRR. D'une manière générale, les créations d'établissements sont positivement corrélées spatialement. Tenir compte de ces effets ne permet pas en revanche de statuer pour un effet significativement différent de la politique sur la création d'établissements dans les communes ZRR. Négliger la présence de ces interdépendances par une modélisation reposant sur les hypothèses CIA et SUTVA aurait au contraire conduit à estimer un effet positif, significatif et biaisé.

Cette contribution rend compte de certaines modalités de prise en compte des interdépendances unissant les comportements des individus. Nous avons proposé un cadre géographique pour approximer ces interdépendances. Ce cadre géographique peut déjà être diversement appréhendé en considérant différentes structures d'interactions et ainsi livrer des analyses de robustesse à nos résultats. D'autres vecteurs sont bien sûr envisageables pour apprécier de manière plus concrète comment ces interactions se nouent, se renforcent ou s'affaiblissent. Nous pouvons par exemple, dans le cas de notre illustration empirique, considérer que les structures intercommunales, les territoires de projets... sont des « lieux » de rencontre entre les acteurs communaux qui rendent compte des réseaux qui les unissent. Il serait par ailleurs possible de définir des structures d'interactions différentes pour le processus de sélection et les effets du traitement. Cette possibilité est appropriée par exemple lorsque la sphère de décision est différente de la sphère d'activation comme dans le cas des politiques d'emplois. Les effets de voisinage sont associés aux lieux d'habitation et aux lieux d'emploi et ceux-ci peuvent être profondément différents. Le domaine de validité des hypothèses permettant de contrôler les interdépendances peut encore faire l'objet d'autres extensions. En effet, la mise en œuvre d'une politique ne conduit pas tous les individus à en bénéficier au même moment. La participation peut se faire au cours du temps. Dans ce cas, tenir compte des interactions devient un élément fort de l'évaluation de la politique, car la participation des nouveaux entrants et les performances de la politique sont potentiellement très dépendantes des premiers bénéficiaires. Ainsi, l'estimation de l'effet de traitement de la politique en prenant en compte à la fois l'aspect dynamique du traitement (Heckman et Navarro, 2007) et les effets d'interaction entre les individus pourrait faire l'objet de recherche future avec comme champs d'application la politique des zones franches urbaines (ZFU).

Annexe 1. La classification en ZRR

Une commune est classée ZRR si elle satisfait les critères suivants (DANIEL et *al.*, 2010) :

- *Un critère démographique* évalué à l'échelle soit des cantons, soit des arrondissements, soit des EPCI selon les données démographiques du recensement de 2009. La commune est incluse (i) dans un canton ayant une densité de population inférieure ou égale à 31 hab./km², ou bien (ii) dans un arrondissement ayant une densité inférieure ou égale à 33 hab./km², ou bien (iii) au 1er janvier 2005 dans un EPCI à fiscalité propre dont le territoire a une densité inférieure ou égale à 31 hab./km². Si ces densités de population sont inférieures à 5 hab./km², alors les communes sont automatiquement classées en ZRR. En revanche, pour les autres communes, une fois le critère démographique satisfait, la commune doit répondre à au moins un des trois critères socio-économiques suivants.
- *Les critères socio-économiques* sont évalués également au niveau des cantons ou des arrondissements. La commune doit avoir connu (i) une baisse de sa population, ou bien (ii) une baisse de sa population active, ou bien (iii) avoir un taux de population active agricole supérieur au double de la moyenne nationale (c'est-à-dire supérieur à 8,3 % de l'ensemble de la population active).
- *Un critère institutionnel* : appartenir à un EPCI à fiscalité propre (à compter de 2005).

Annexe 2. Description et sources de données

Les données utilisées, accessibles en ligne, sont issues d'un ensemble de bases de données produites principalement par l'Insee et la Datar.

- Le classement de l'ensemble des communes provient du *fichier historique des communes classées en ZRR* de l'Observatoire des territoires (Datar). On trouve dans ce fichier, le classement des communes en ZRR depuis 1995 jusqu'à 2013. Pour les besoins de l'analyse nous avons utilisé les données de 2010 selon le Code Officiel de la Géographie COG de septembre 2010 (COG0910). <http://www.observatoire-des-territoires.gouv.fr/observatoire-des-territoires/fr/historique-des-communes-class-es-en-zrr> (page consultée le 28/03/2014).
- Les données relatives à la création d'établissements en 2011, sont extraites du fichier « *Démographie des entreprises 2006-2011* » de l'Insee. Ces données sont disponibles à l'échelle de la commune et selon le type d'activités en 4 postes (industrie, construction, commerce-transport-service, administration publique-enseignement-santé-action sociale). Le champ d'activité agriculture est exclu de l'analyse.
http://www.insee.fr/fr/themes/detail.asp?reg_id=99&etref_id=base-cc-demo-entreprises (Page consultée le 28/03/2014)
- Les données sur le zonage en aires urbaines proviennent de la base communale des aires urbaines 2010 de l'Insee. Compte tenu de la particularité des ZRR, nous avons retenu quatre types de zones au sens du *nouveau zonage en aires urbaines de 2010* (Insee, ZAU 2010), en excluant les communes appartenant aux zones relatives aux *grandes aires urbaines*. Les communes retenues appartiennent aux : (i) "*moyennes aires*" définies comme étant « ensemble de communes, d'un seul tenant et sans enclave, constitué par un pôle (unité urbaine) de 5 000 à 10 000 emplois, et par des communes rurales ou unités urbaines dont au moins 40 % de la population résidente ayant un emploi travaille dans le pôle ou dans des communes attirées par celui-ci ». (ii) "*petites aires*", « ensemble de communes, d'un seul tenant et sans enclave, constitué par un pôle (unité urbaine) de 1 500 à 5 000

emplois, et par des communes rurales ou unités urbaines dont au moins 40 % de la population résidente ayant un emploi travaille dans le pôle ou dans des communes attirées par celui-ci ». (iii) les autres communes multi-polarisées et (iv) les communes isolées, hors influence des pôles.

http://www.insee.fr/fr/methodes/default.asp?page=zonages/aires_urbaines.htm.
(Page consultée le 24/03/2014)

- Les données « *Typologie thématique du champ "Paysages"* » proviennent de la Datar. Un zonage des communes de la métropole a été fait en fonction de trois critères : composition de l'occupation du sol, configuration de l'occupation du sol et relief (pente moyenne, dénivellation). On retrouve au total dix groupes ou zones : Rural artificialisé : présence forte du bâti en paysage diversifié (cultures, forêts, vignobles); Collines et vallées, paysages diversifiés (bâti, cultures, prairies, forêts); Paysages artificialisés en contexte semi-naturel ou humide, notamment littoraux; Collines et vallées en prairies, peu artificialisé; Paysages de cultures avec forte marque du bâti; Relief marqué, paysages diversifiés : forêts, prairies, milieux semi-naturels; Plaines et plateaux et grands openfields; Montagne, très forts reliefs et végétation semi-naturelle; Plaines et plateaux, campagne diversifiée avec bâti marqué; Plaines et plateaux, prairies et forêts. <http://www.observatoire-des-territoires.gouv.fr/observatoire-des-territoires/fr/typologie-thematique-du-champ-paysages>. (Page consultée le 24/03/2014)
- Les données socioéconomiques proviennent du fichier *Résultats du recensement de la population 2010* de l'Insee et sont disponibles au niveau de la commune. Elles regroupent des données relatives à la population des communes, la structure du marché de l'emploi (population active, population active en activité et population active au chômage), emploi au lieu de travail (statut, temps de travail, catégories socioprofessionnelles et secteurs d'activité économique).
<http://www.insee.fr/fr/bases-de-donnees/default.asp?page=recensement/resultats/2010/donnees-detaillees-recensement-2010.htm> (page consultée, le 21/03/2014).

**Chapitre 5 Bounding treatment effect and interactions:
A spatial economics approach applied to the French
“Zones Urbaines Sensibles” program**

Introduction

The treatment effect estimation methods used extensively in the literature on evaluation are based on the Rubin causal model (RCM) with strong assumptions that are questionable in the context of nonexperimental policies. As indicated by Lechner (1999 p. 2): "*If assignment to the treatment and control group is not random, knowledge of the assignment mechanism is necessary to adjust these estimates of the mean accordingly and avoid selection bias*". Since in practice, incomplete information is more the rule, it is impossible to control for assignment to treatment, and additional information and processes are required in order to control for potential selection bias. Another solution is to adopt experimental processes, but they are costly to generalize and in some case studies perhaps impossible to perform. Accordingly, "*more or less serious doubts about the validity of the chosen identifying assumptions plague almost every empirical evaluation study based on nonexperimental data*" (Lechner, 1999, p. 2).

Manski's method known as *nonparametric bounds* (1989, 1990) takes a different approach: given the incomplete information sets concerning the participation process, the true value of parameters cannot be accurately evaluated. It is, however, possible to give a range of estimations for the treatment effects, i.e. to define bounds, without requiring strong assumptions about the selection mechanisms. This advantage opens up the analysis of treatment effects to several extensions and applications in socio-economics. For example, Manski et al. (1992) and Horowitz and Manski (1999) estimate the effect of family structure on high school graduation. They use a bootstrap procedure on their data to estimate the bounds of the treatment effect. Huang et al. (2012) analyze the effect of educational attainment on social capital by means of a multi-level education model. In their case, the conventional Instrumental Variables (IV) method is not entirely efficient and strong functional form assumptions are required. The authors use a nonparametric strategy to overcome this problem. Exploring the effect of return to school on earnings in post-Apartheid South Africa, Mariotti and Meinecke (2009) suggest that the large spread of results estimated in earlier research may arise because the data are sensitive to the parametric methodology used in the studies. In their study of the effect of return to education for African males in South Africa, they extend the nonparametric bounds to the problem of self-selection because the constraints on freely selected behavior in education may not be ignored in South Africa. Recently, De Haan (2014) argues that the evaluation of education resources on pupil

outcomes, while an important and serious matter, is rather controversial. So, many varied policies are applied that preclude homogeneous results even when strong assumptions and recent econometric methods are used and the identification problem remains difficult to deal with. On the contrary, nonparametric bounds analysis appears more convenient for De Haan (2014), even in the presence of heterogeneity in the impacts of the “Learning Support” program she evaluates. Although a range instead of a point value estimate is produced by the nonparametric bounds approach, it is informative in that case and reveals a significant positive impact of additional resources on pupil achievement.

Extensions to nonparametric bounds estimation methods have not yet addressed interdependencies among people although social studies widely stress the impact of neighborhood effects on behaviors. People are still assumed to be isolated from interference effects. However, interaction effects are commonly observed and often permeate selection steps and policy outcomes. For example, people who might be eligible for assistance under a given policy may not come forward because they do not want their neighbors to know about the difficulties they face. In the context of a regional vaccination program, the more people are vaccinated, the better the health of the whole population is protected. An educational program helping just some students in a class can improve the result of the rest of the class through peer effects. In labor markets, training programs for the unemployed may promote better matches with jobs and so encourage more unemployed people to take the training, which will reduce their chances of getting back latter into work.

Recently, some authors have pointed out that interdependencies between individuals mean the evaluation of treatment effects should be reconsidered (Flores-Lagunes and Schnier, 2012; Ba and Baumont 2015). This implies reconsidering the validity of the Stable Unit Treatment Value Assumption (SUTVA), which assumes that the treatment does not produce externalities on untreated people/other treated people. But when such interactions do exist but are not addressed, the estimates may be affected by the endogeneity problem, which can take on either of two forms: omitted variable bias or spatial autocorrelation bias. Therefore, the economic implications drawn from such results are also biased.

With the nonparametric bounded treatment approach, it can be argued that interaction effects are one source of unavailable information, precluding the use of standard evaluation methods. This is true and it reinforces the appropriateness of nonparametric bounds estimates. However, in the spirit of nonparametric approaches, modeling interaction patterns may be one

way to analyze whether and how human behaviors contribute to treatment outcomes even in their bounded values.

In a general setting, these interaction effects have been classified by Manski (1993) into three types: *endogenous*, *exogenous*, and *correlated*. They question why individuals belonging to the same group tend to act similarly. This classification has been extensively applied to neighborhood effects with the underlying assumption that to belong to a group could be considered similar to living in the same neighborhood. *Endogenous effects* refer to interdependencies between individual decisions: the behavior of one individual directly influences the behavior of other individuals in the group (or living in the same neighborhood). *Contextual* or *exogenous effects* occur if the behavior of one individual depends on the exogenous characteristics of a group of people (or of neighbors). *Correlated effects* occur because the individuals belonging to a group (or living in the same neighborhood) tend to share similar institutional exposures.

Beyond the simple assimilation between the fact that people who belong to a given social group tend to live in the same place (neighborhood), the assimilation between social interactions and neighborhood effects raises three challenges. The first is that social interaction patterns are in some sense geographically defined. The second is that interdependencies may be formalized through spatial effects and especially in the form of spatial autocorrelation which measures the coincidence of value similarity with locational similarity (Anselin 2001). More precisely, social interactions and neighborhood interactions are similar if people who share the same attributes or behave similarly tend to cluster in the same neighborhood. Third, spatial analysis methods can be implemented that open up to a more general framework to include interdependencies in treatment effect estimations. In fact, modeling spatial patterns of interactions allows some overlap of patterns between neighborhoods and groups. With neighborhood effects *stricto sensu*, the neighborhood is the place, the block, or the district, where the social group lives. It is generally bounded by physical or social frontiers, and people interact with each other within the neighborhood. One challenge is to deal with the frontier effect. In spatial data analysis, the neighborhood is the set of people with whom a given individual interacts. So some individuals may have a larger set of neighbors than others and neighborhoods may overlap. Moreover, no restriction is made on neighborhood size or the number of neighbors because the only assumption to be complied

with is that interaction between two people does not decline with their degree of similarity. Accordingly two individuals belonging to two separate social groups but living in the same building may interact more than two individuals of the same social condition but living 10 kilometers apart. Of course, interactions between people of similar social conditions and living in the same neighborhood may be greater. For the treatment effect evaluation, the interactions between two pupils in the same classroom, one receiving some educational assistance but the other not, illustrates the former situation.

Given this context, the aim of this chapter is to provide a technical framework based on spatial econometrics methods to take account of interaction effects when estimating treatment effect bounds.

In the following parts, we first present the nonparametric bounds method in a-spatial and spatial frameworks (5.1). More precisely, we suppose that the outcome observed for one individual regardless of status (treated or not treated) depends on the individual’s own characteristics but also on the neighbors’ observed outcomes. This hypothesis holds through different diffusion mechanisms (learning, competition, imitating, coercion)⁶⁹ (Shipan and Volden 2008) meaning that one treatment applied to a group of individuals can affect the outcome for the rest of the population. We model the spatial extension of the parameters of interest: the average treatment effect and the average treatment effect on the treated in a-spatial and spatial frameworks. The remaining parts of the chapter develop the empirical illustration. Social interactions have been largely documented in urban renewal policies especially through the impacts of neighborhood effects on economic outcomes or social outcomes of people living in underprivileged districts. We choose to apply nonparametric bounds estimates to the French “Zones Urbaines Sensibles” program. We analyze the effect of this place-based policy on the unemployment rate at the neighborhood level. The ZUS policy, the data, and the variables are presented first (5.2). Treatment effect bounds are estimated and the results are presented in the final section (5.3). We estimate two types of models comparing the traditional (without interdependencies) and the spatial specification, each of them being estimated under the no-assumption hypothesis and the monotone treatment assumption. The chapter concludes with suggestions for further research.

⁶⁹ See Chapter 1.

5.1 Spatial modeling of nonparametric bounds on treatment effects

5.1.1 The identification of outcomes in the presence of interactions

In a population of N members, let us consider an individual i characterized by a set of variables (y_1, y_0, d, x) where x denotes the vector of the individual's characteristics, d is a binary variable that indicates whether the individual i is treated or not, y_1 is the outcome observed if individual i is treated and y_0 is the potential outcome if individual i is untreated.

For each individual, we observe:

$$\text{And} \quad y_1 = x' \beta_1 + \varepsilon_1 \quad (5-1)$$

$$y_0 = x' \beta_0 + \varepsilon_0 \quad (5-2)$$

ε_1 and ε_0 are i.i.d. error terms and β_1 and β_0 are the parameters associated with the covariate x in each potential outcome equation.

In matrix form, the specifications (5-1) and (5-2) can be written as follows:

$$\text{And} \quad Y_1 = X\beta_1 + \varepsilon_1 \quad (5-3)$$

$$Y_0 = X\beta_0 + \varepsilon_0 \quad (5-4)$$

Y represents an $(N \times 1)$ vector of the observed outcomes, Y_1 is the vector of the observed outcomes when the individuals are treated, and Y_0 is the vector of the observed outcomes when individuals are untreated. X denotes a $(N \times p)$ matrix of the p explanatory variables. β denotes a $(p \times 1)$ vector of the associated parameters. ε_1 and ε_0 are the $(N \times 1)$ vectors of the error terms that are independent and identically distributed following a normal distribution with mean 0 and variance σ^2 .

Taking into account interactions between individuals in the estimation of treatment effects requires some assumptions to describe the interaction patterns and the treatment

process. More precisely, we suppose some spillover effects from the treatment. This means that the policy affects not only the outcomes of the treated but also the outcomes of the people located in the neighborhoods, too, whatever their status: treated or not treated. This first assumption makes social overlapping possible. Second, we suppose that these spillovers are generated by a spatially dependent process which is formalized by a spatial autoregressive specification (SAR model) of equations (5–3) and (5–4). This means that the potential outcome observed for each individual depends on the potential outcome observed in their neighborhood.

The observed outcome takes the form:

$$Y = \rho WY + \beta X + \varepsilon \quad (5-5)$$

Where Y , X , and ε are defined as before.

W is the $(N \times N)$ spatial weights matrix describing the interaction patterns. It is exogenously defined to reflect the structure of interdependencies among all members of the population. More precisely, for each individual i , the term w_{ij} indicates the neighbor j she interacts with ($w_{ij} \neq 0$) or not ($w_{ij} = 0$). The larger w_{ij} is, the greater is the interaction between i and j . It is assumed that the more similar the individuals are, the more they interact. $w_{ii} = 0$ by assumption.

ρ is the spatial parameter to be estimated and it gives the spillover effects resulting from the interaction pattern across the population.

The spatial autoregressive model allows individual neighborhoods to overlap because the sets of neighbors of each individual i are not necessary identical. Moreover, neighborhood overlapping is also reflected by the spatial multiplier effects. In fact, the spatial autoregressive equation (5–5) may be written as follows, if the matrix $(I - \rho W)$ is not singular:

$$Y = (I - \rho W)^{-1}\beta X + (I - \rho W)^{-1}\varepsilon \quad (5-6)$$

Equation (5–6) shows that the potential outcome observed for an individual is not only affected by that individual’s own characteristics and random shocks but it is also affected by the characteristics and random shocks of the other individuals and their neighborhoods through the spatial multiplier $(I - \rho W)^{-1}$.

The value of the spatial parameter gives three results regarding the spillovers and interdependencies between people.

Proposition 1: if $\rho = 0$, there is no interaction between individuals. The treatment received by one individual affects only her own outcome. Then the bounds of the interest parameters can be estimated with Manski's standard approach ignoring the interactions.

Proposition 2: if $\rho > 0$, there are positive interactions between individuals. The outcome observed for one individual tends to be similar to the outcomes observed for her neighborhood. Social interactions support a reinforcement effect (Manski, 2013). Moreover, the treatment received by one individual affects not only her own outcome but also the outcome of the neighbors in the whole population due to the spatial multiplier effect.

Proposition 3: if $\rho < 0$, there are negative interactions among individuals. The outcome of one individual and the outcomes observed in her neighborhood tend to be different. Social interactions tend to support an opposing effect (Manski, 2013) and the treatment received by one individual affects not only her own outcome but also the outcome of the neighbors in the whole population due to the spatial multiplier effect.

The impacts of interactions on the treatment effects are then given by the estimates of the spatial specification of the parameters of interest.

5.2 Interactions and identification of treatment effects

The main parameters of interest widely used to measure the treatment effect in the literature are: (i) the average treatment effect (ATE) and (ii) the average treatment effect on the treated (ATT).

5.2.1 Identification of the Average Treatment Effect

Given a set of observable characteristics, ATE is defined as the difference between average observed outcomes of the treatment group Y_1 and the average observed outcomes of the nontreatment group Y_0 . The latter is also referred to as the control group.

$$\text{ATE} = E(Y_1|X) - E(Y_0|X) \quad (5-7)$$

Using the law of iterated expectations (Manski, 1989), we obtain for each group:

$$E(Y_1|X) = E(Y_1|X, D = 1)P(D = 1|X) + E(Y_1|X, D = 0)P(D = 0|X) \quad (5-8)$$

$$E(Y_0|X) = E(Y_0|X, D = 0)P(D = 0|X) + E(Y_0|X, D = 1)P(D = 1|X) \quad (5-9)$$

D is a binary assignment indicator that indicates whether the individual receives the treatment ($D=1$) or not ($D = 0$). Then the terms $P(D = k|X)$ for $k = 1$ or 0 are the probabilities that one individual receives or does not receive the treatment given his/her characteristics.

The parameter of interest ATE then becomes:

$$\begin{aligned} \text{ATE} &= E(Y_1|X) - E(Y_0|X) \\ &= [E(Y_1|X, D = 1)P(D = 1|X) + E(Y_1|X, D = 0)P(D = 0|X)] \\ &\quad - [E(Y_0|X, D = 0)P(D = 0|X) \\ &\quad + E(Y_0|X, D = 1)P(D = 1|X)] \\ &= [E(Y_1|X, D = 1) - E(Y_0|X, D = 1)]P(D = 1|X) \\ &\quad + [E(Y_1|X, D = 0) - E(Y_0|X, D = 0)]P(D = 0|X) \\ &= [E(Y_1|X, D = 1) - E(Y_0|X, D = 1)]P(D = 1|X) \\ &\quad + [E(Y_1|X, D = 0) - E(Y_0|X, D = 0)](1 - P(D = 1|X)) \end{aligned} \quad (5-10)$$

Once everyone has received or not received the treatment, it is possible to observe the terms $E(Y_1|X, D = 1)$, $P(D = 1|X)$, $E(Y_0|X, D = 0)$ and $P(D = 0|X)$, but it is impossible to

observe the terms $E(Y_1|X, D = 0)$ and $E(Y_0|X, D = 1)$. One and the same individual cannot be in both states at once: treated and not treated.

The main idea in Manski's approach is to consider that the outcome Y_1 is bounded by Y_1^l and Y_1^u , considered respectively as the lower and upper bounds of Y_1 . In other words Y_1^l is the smallest value that Y_1 can take whereas Y_1^u is the largest value that Y_1 can take. We have the same considerations for the observed outcome Y_0 . It is always possible to observe such information in the population and substitute it for the unobservable terms. These values are considered in a first step to define the no-assumption case (Manski, 1990) which means that we consider what we can learn from the data alone.

Under the no-assumption hypothesis, for the treatment group, the term $E(Y_1|X, D = 0)$ can then take two values Y_1^l and Y_1^u and we replace $E(Y_1|X, D = 0)$ in equation (5-8), by Y_1^l and Y_1^u .

$$\begin{aligned} E(Y_1|X, D = 1)P(D = 1|X) + Y_1^l P(D = 0|X) \\ \leq E(Y_1|X) \leq \\ E(Y_1|X, D = 1)P(D = 1|X) + Y_1^u P(D = 0|X) \end{aligned}$$

If we denote B_1^l and B_1^u the left and right hand sides of the intervals under the no-assumption hypothesis, it can be rewritten as:

$$B_1^l \leq E(Y_1|X) \leq B_1^u \quad (5-11)$$

With:

$$\begin{aligned} B_1^l &= E(Y_1|X, D = 1)P(D = 1|X) + Y_1^l P(D = 0|X) \\ &= E(Y_1|X, D = 1)P(D = 1|X) + Y_1^l (1 - P(D = 1|X)) \end{aligned}$$

And:

$$\begin{aligned} B_1^u &= E(Y_1|X, D = 1)P(D = 1|X) + Y_1^u P(D = 0|X) \\ &= E(Y_1|X, D = 1)P(D = 1|X) + Y_1^u (1 - P(D = 1|X)) \end{aligned}$$

Similarly, for the control group, under the hypothesis of no-assumption bounds, the term $E(Y_0|X, D = 1)$ can then take two values Y_0^l and Y_0^u and we replace $E(Y_0|X, D = 1)$ in equation (5–9), by Y_0^l and Y_0^u . Using B_0^l and B_0^u for the left and right hand sides of the intervals under the no-assumption hypothesis, we obtain:

$$B_0^l \leq E(Y_0|X) \leq B_0^u \quad (5-12)$$

With:

$$\begin{aligned} B_0^l &= E(Y_0|X, D = 0)P(D = 0|X) + Y_0^lP(D = 1|X) \\ &= E(Y_0|X, D = 0)(1 - P(D = 1|X)) + Y_0^lP(D = 1|X) \end{aligned}$$

And:

$$\begin{aligned} B_0^u &= E(Y_0|X, D = 0)P(D = 0|X) + Y_0^uP(D = 1|X) \\ &= E(Y_0|X, D = 0)(1 - P(D = 1|X)) + Y_0^uP(D = 1|X) \end{aligned}$$

Equations (5–11) and (5–12) mean that the expected outcomes Y_k (for the treated group if $k=1$ and the control group if $k=0$) conditional on $(X = x)$ are bounded respectively by the lower bound and upper bound of the expected outcomes.

It follows that the average treatment effect ATE (equation (5–10) is also bounded by:

$$\begin{aligned} [B_1^l - B_0^u] \leq ATE \leq [B_1^u - B_0^l] \\ B_{ATE}^l \leq ATE \leq B_{ATE}^u \end{aligned} \quad (5-13)$$

After rearranging the extensive form of the lower bound B_{ATE}^l we obtain:

$$\begin{aligned} B_{ATE}^l &= \{E(Y_1|X, D = 1) - Y_0^u\}P(D = 1|X) \\ &\quad + \{Y_1^l - E(Y_0|X, D = 0)\}(1 - P(D = 1|X)) \end{aligned} \quad (5-14)$$

The upper bound B_{ATE}^u is respectively expressed as:

$$\begin{aligned} B_{ATE}^u &= \{E(Y_1|X, D = 1) - Y_0^l\}P(D = 1|X) \\ &\quad + \{Y_1^u - E(Y_0|X, D = 0)\}(1 - P(D = 1|X)) \end{aligned} \quad (5-15)$$

Considering the spatial specification of outcomes as in equation (5–6), we obtain the expected values of outcomes including the interactions between people.

The two spatial bounds are denoted respectively by SB_{ATE}^l and SB_{ATE}^u . Table 5-1 gives the expressions of the estimates in the a-spatial and spatial cases. The difference between the two estimates is the introduction of interaction effects in the estimation of bounds with the matrix $(I - \rho_1 W)^{-1}$ and $(I - \rho_0 W)^{-1}$.

Table 5-1: The estimated bounds of ATE

a-spatial framework	
B_{ATE}^l	$\{x\beta_1 - Y_0^u\}P(D = 1 X) + \{Y_1^l - x\beta_0\}(1 - P(D = 1 X))$
B_{ATE}^u	$\{x\beta_1 - Y_0^l\}P(D = 1 X) + \{Y_1^u - x\beta_0\}(1 - P(D = 1 X))$
Spatial framework	
SB_{ATE}^l	$\{(I - \rho_1 W)^{-1}x\beta_1 - Y_0^u\}P(D = 1 X) + \{Y_1^l - (I - \rho_0 W)^{-1}x\beta_0\}(1 - P(D = 1 X))$
SB_{ATE}^u	$\{(I - \rho_1 W)^{-1}x\beta_1 - Y_0^l\}P(D = 1 X) + \{Y_1^u - (I - \rho_0 W)^{-1}x\beta_0\}(1 - P(D = 1 X))$

5.2.2 Identification of the Average Treatment Effect of the Treated

The second parameter of interest, ATT, is defined for the treatment group and is the difference between the average observed outcomes under treatment and the average observed outcomes without treatment.

$$ATT = [E(Y_1|X, D = 1) - E(Y_0|X, D = 1)] \quad (5-16)$$

The ATT is bounded by:

$$B_{ATT}^l \leq ATT \leq B_{ATT}^u$$

With:

$$B_{ATT}^l = \{E(Y_1|X, D = 1) - Y_0^u\}$$

And:

$$B_{ATT}^u = \{E(Y_1|X, D = 1) - Y_0^l\}$$

When considering the spatial case, we obtain the spatial upper bound SB_{ATT}^u and the spatial lower bound SB_{ATT}^l . The estimated values are displayed in Table 5-2 and the difference between the standard situation and the case where people interact with each other is the presence of an interaction effect in the estimation of bounds with the matrix $(I - \rho_1 W)^{-1}$.

Table 5-2: The estimated bounds of ATT

	a-spatial framework
B_{ATT}^l	$X\beta_1 - Y_0^u$
B_{ATT}^u	$X\beta_1 - Y_0^l$
	Spatial framework
SB_{ATT}^l	$(I - \rho_1 W)^{-1}X\beta_1 - Y_0^u$
SB_{ATT}^u	$(I - \rho_1 W)^{-1}X\beta_1 - Y_0^l$

Once the bounds of the treatment effects have been defined, their usefulness for policy evaluation has to be discussed.

5.2.3 Assumptions about the range of bounds

When there is no assumption about the selection process, one has to wonder about the width of the estimated bounds of the treatment effect, which can be very wide indeed. The width can include the null and the sign of the treatment effect is not identified. The estimated bounds are therefore not helpful to know the true effect of the treatment. This is the *worst-case*⁷⁰ also called the *no assumption bound* in the literature. Lechner (1999) compares different assumptions to improve the performance of the nonparametric bounds method: i.e. tighten the bounds and narrow the interval. To overcome the worst-case problem, Manski (1990) and Manski and Pepper (2000) build a contextual framework to tighten the estimated bounds based on two assumptions, which together improve the precision of the bounds and give more indications about the sign of the treatment effect. The first assumption – Monotone

⁷⁰ Note that in some cases, the worst case is not useful because Y_1^l and Y^u cannot be identified, i.e. they may be infinite. In other cases, when the outcome is binary, i.e. Y_1 and $Y_0 \in [0,1]$, then the range of bounds is equal to 1.

Treatment Selection assumption (MTS) – concerns the selection process. The second assumption – Monotone Treatment Response (MTR) – concerns the observed outcomes.

The MTS assumption is centered on the fact that some individuals decide to join the treatment or not for different reasons that are unobserved in many cases. So if we compare two groups, those who actually received a low treatment level and those who actually received a higher level of treatment, then, the average outcomes would be higher in the group that actually received the higher level of treatment than they would be in the group that actually received the lower level.

Following Manski (1990) and Manski and Pepper (2000), the MTS assumption is expressed as follows:

Let T be an ordered set of treatment levels, t_0 and t_1 two elements of T , and z is the realized treatment $z \in T$.

$$t_1 \geq t_0 \Rightarrow E[y(t)|z = t_1] \geq E[y(t)|z = t_0]$$

In an application to returns to schooling, Manski and Pepper (2000, p. 1001) emphasize that: “*the MTS assumption asserts that persons who select higher levels of schooling have weakly higher mean wage functions than do those who select lower levels of schooling*”. This assumption allows us to control the sample selection problems such as selection bias (Mariotti and Meinecke 2009).

The MTR assumption is centered on the fact that the treatment is supposed to increase the observed outcome and then strongly predetermines the sign of the treatment. In other words, after the treatment, it is expected that the observed outcome of the treated individuals would be greater than the observed outcome of the untreated individuals. Moreover, as emphasized by VanderWeele *et al.* (2014, p. 3): “*The MTR assumption is simply that for each individual an increase in the treatment variable A will increase or leave unchanged the outcome that would be observed*”.

Following Manski (1990) and Manski and Pepper (2000), the MTR assumption is formalized as follows, for each individual i of population I .

Let T be an ordered set, t_0 and t_1 two elements of T .

$$t_1 \geq t_0 \implies y_i(t_1) \geq y_i(t_0)$$

In their application to the returns to schooling, Manski and Pepper (2000, p. 999) indicate that: “*The MTR assumption asserts that, ceteris paribus, wage rises as a function of conjectured years of schooling*”.

Under these assumptions (MTR and MTS), the sharp bound of $E(Y_1|X)$ and $E(Y_0|X)$ is expressed as follows:

$$\begin{aligned} & \sup \{E(Y_1|X, D = 1)P(D = 1|X) + Y_1^l (1 - P(D = 1|X))\} \\ & \leq \mathbf{E}(Y_1|X) \leq \end{aligned} \tag{5-17}$$

$$\inf \{E(Y_1|X, D = 1)P(D = 1|X) + Y_1^u (1 - P(D = 1|X))\}$$

$$\begin{aligned} & \sup \{E(Y_0|X, D = 0)(1 - P(D = 1|X)) + Y_0^l P(D = 1|X)\} \\ & \leq \mathbf{E}(Y_1|X) \leq \end{aligned} \tag{5-18}$$

$$\inf \{E(Y_0|X, D = 0)(1 - P(D = 1|X)) + Y_0^u P(D = 1|X)\}$$

From equation (5-17) and (5-18), the ATE is bounded by:

$$\begin{aligned} & \left[\sup \{E(Y_1|X, D = 1)P(D = 1|X) + Y_1^l (1 - P(D = 1|X))\} \right. \\ & \quad \left. - \inf \{E(Y_0|X, D = 0)(1 - P(D = 1|X)) + Y_0^u P(D = 1|X)\} \right] \\ & \leq \mathbf{ATE} \leq \end{aligned} \tag{5-19}$$

$$\begin{aligned} & \left[\inf \{E(Y_1|X, D = 1)P(D = 1|X) + Y_1^u (1 - P(D = 1|X))\} \right. \\ & \quad \left. - \sup \{E(Y_0|X, D = 0)(1 - P(D = 1|X)) + Y_0^l P(D = 1|X)\} \right] \end{aligned}$$

To estimate the bounds, either under the worst case or with the additional assumption case, several approaches are proposed. Many assumptions about the selection process and the outcome have to be specified with regard to the parametric methods for estimating the terms defining the bounds (equations (5–3) and (5–4)). For example, Huang et al. (2012) justify their choice of nonparametric estimators to control for endogeneity of education on social capital: *“the OLS model may produce a biased and inconsistent estimate of the educational effect when the choice of educational attainment and social capital are simultaneously influenced by unobserved heterogeneity specific to the individuals”* (Huang et al. (2012, p. 1024). Some authors advocate the Instrumental Variable Method (IVM) when one suspects that the independence assumption does not hold but, once again, some specific assumptions have to be postulated. One of the main limits is the difficulty in finding good instruments, i.e. instrumental variables which affect the treatment directly and the potential outcome through the treatment indirectly. Gerfin and Schellhorn (2006, p. 9) underline that *“necessary exclusion restrictions for instrumental variables and selection models are often difficult to justify”*. Manski and Nagin (1998) demonstrate that the IVM in the parametric framework presents some limits because it assumes that the treatment effect is constant in the population. In addition, they show that few explanatory variables affect only the selection process and not the outcome results. Ravallion (2005, p. 59) points out: that *“In practice, IVE⁷¹ sometimes gives seemingly implausible impact estimates (either too small or too large)”*.

On the other hand, the nonparametric methods are attractive for estimating bounds because they do not *“rely on delicate functional and distributional assumptions”⁷²* while *“some smoothing methods (kernels, nearest neighbour, series estimation...) could be used, in principle, to compute the bounds to improve efficiency”⁷³*. Even so, these nonparametric methods require large sample sizes to perform acceptably; additionally they must be driven depending on the type of data; and they appear unsuitable in the case of discrete data (Lechner, 1999).

The nonparametric bounds approach has received much attention in the literature on evaluation. We suggest extending this method to take into account interdependencies among people. In fact, this issue has already been addressed in one case. Considering that spillovers

⁷¹ IVE: Instrumental Variable Estimator.

⁷² Huang et al. (2012, p. 1022)

⁷³ Lechner (1999, p. 9)

may occur among people, De Haan (2014) supposes an additional hypothesis she calls the Monotone Spillover assumption (MTS). In her case study about returns on education, De Haan (2014, p. 17) emphasizes that: “*The monotone spillover assumption states that the mean examination results of the ineligible pupils ($z^{74} = 0$) in the case where no pupil would receive Learning Support, would not be higher than their current mean examination results. Similarly, MSP assumes that the mean examination results of pupils eligible for Learning Support ($z = 1$) in the case where the schools would receive additional funding for all pupils, would not be lower than their current mean examination results, which are realized in a setting in which schools receive additional funding for only part of the pupils*”. However, she does not indicate how to integrate these spillovers into the estimation of the treatment effect and she only considers positive externalities from the treatment.

Our methodology differs since we propose a general framework to include all types of spillover effects in the nonparametric bounds approach.

5.3 The French *Zones Urbaines Sensibles* (ZUS) program

The case study is the impact of the French “*Zones Urbaines Sensibles*” (ZUS) program on the unemployment rate of people living in ZUSs in the city of Paris or in its immediate suburbs. We briefly present the ZUS Program and the study area (5.2.1) before focusing on the data (5.2.2) and the identification of the sample of ZUSs (5.2.3). The treatment group and the control group are obtained with the propensity score matching method applied to a logit regression.

5.3.1 The ZUS program

In December 1996, the French government launched a major policy known as the “Urban Revival Act” to revitalize those deprived urban districts called *Zone Urbaines Sensibles* (Sensitive Urban Zones) that have to face numerous difficulties in terms of social, economic, and living conditions. The ZUS program is an area-based policy targeting certain geographical zones and is similar to New Deal Community programs in Great Britain for example. The French national institute of statistics and economics studies (Insee) indicates that ZUSs are “*infra-urban territories defined by the authorities as being priority targets for*

⁷⁴ Compared to the previous formulation we have written $z = t$ in the MTS and MTR assumptions.

urban policy, according to local factors relating to the difficulties that the inhabitants of these territories are experiencing”.⁷⁵ The ZUS is “characterized by the presence of large groups of degraded habitat districts and a marked imbalance between housing and employment”.⁷⁶ In short, ZUSs can be defined as being the most distressed zones of a municipality in terms of unemployment, housing, social exclusion, school failure, violence,...

For an area to be classified as a ZUS, four indicators are considered:⁷⁷ the unemployment rate, the proportion of residents with no qualifications, the proportion of residents under the age of 25, and the tax potential of the hosting municipality. In all, 717 ZUSs have been targeted in mainland France, many of them located in large cities and their suburbs. The common goals of the ZUS program are to improve employment, attract firms and economic activities, improve school success, improve inhabitant’s income and living conditions regarding housing, health, security, and so on. The advantages offered by the program are proportional to the territory’s difficulties and some ZUSs benefit from additional public policies: a ZFU “*Zone Franche Urbaine*” is a ZUS confronted with serious economic and social difficulties and will benefit from more assistance and specific fiscal policies especially for enterprises (tax breaks and exemptions). A ZRU (“*Zone de Rénovation Urbaine*”, Urban Renewal Zone) faces serious difficulties, too, but less than a ZFU and will benefit from the same types of fiscal policies as a ZFU.

In short, the ZUS program focuses on the more deprived zones and tries to improve their inhabitants’ living conditions, economic status, and social capital. The main indicators for gaging that impact are job creation, business creation, the refurbishment or replacement of housing, unemployment rate changes, and so on.

5.3.2 The case study: the Paris Area

Our case study focuses on the city of Paris and its suburbs, which cover the first ring around Paris known as the “*Petite Couronne*”. We denote this space as the Paris area (*cf.*

⁷⁵ <http://www.insee.fr/en/methodes/default.asp?page=definitions/zone-urbaine-sensible.htm>

⁷⁶ <http://www.onzus.fr/presentation/les-quartiers-de-la-politique-de-la-ville>

⁷⁷ In 2013, a New Policy for Cities was enacted by the French Government. The selection criteria have changed. Our empirical study does not cover this new program.

Table 5-3 and Map 5-1). It is an urban and very densely populated space, with an average of 8540 inhabitants per square km (min = 5300 for the Val-de-Marne and max = 20775 for the city of Paris). The Paris area is composed of 124 municipalities (or 143 municipalities when we consider the division of the city of Paris into its 20 *arrondissements*). The total area extends over 762.4 km² and had a population of about 6.5 million people in 2006. In all 84 territories have been classified as ZUSs, i.e. more than 53 % of the ZUSs in the Ile-de-France Region (with a total of 157 ZUSs) and more than 10 % of the ZUSs in France. Some 12 % of the population of the Paris area lives in a ZUS.

Map 5-1 : The Paris area: Paris and the *Petite Couronne*

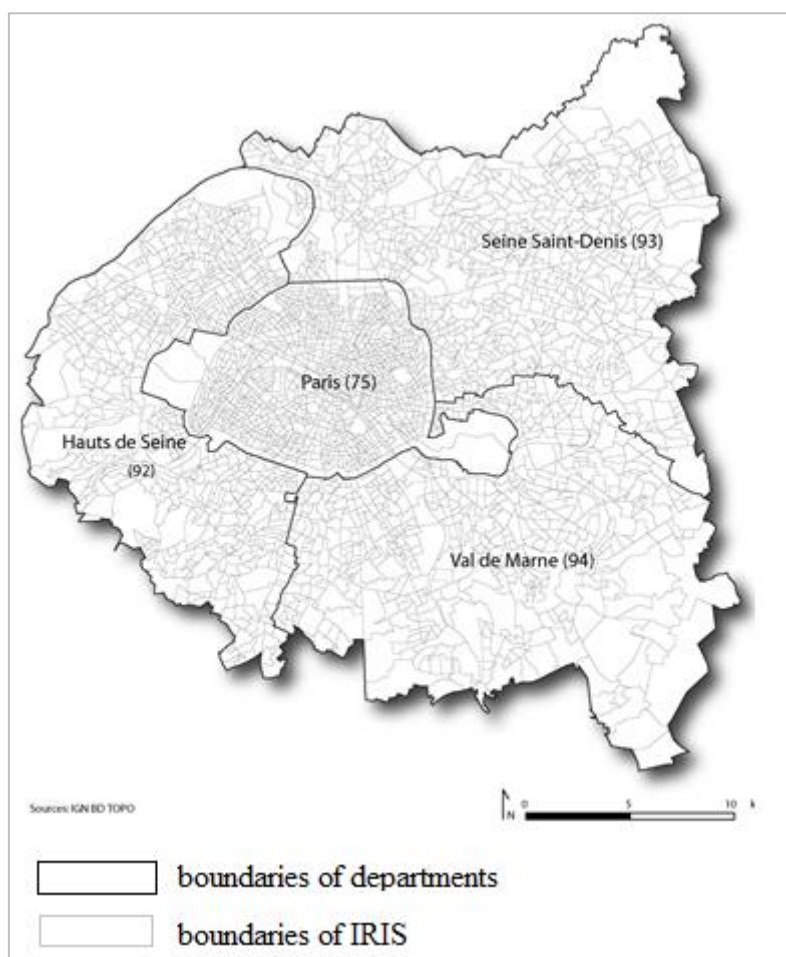


Table 5-3: ZUSs in Paris and its surrounding departments

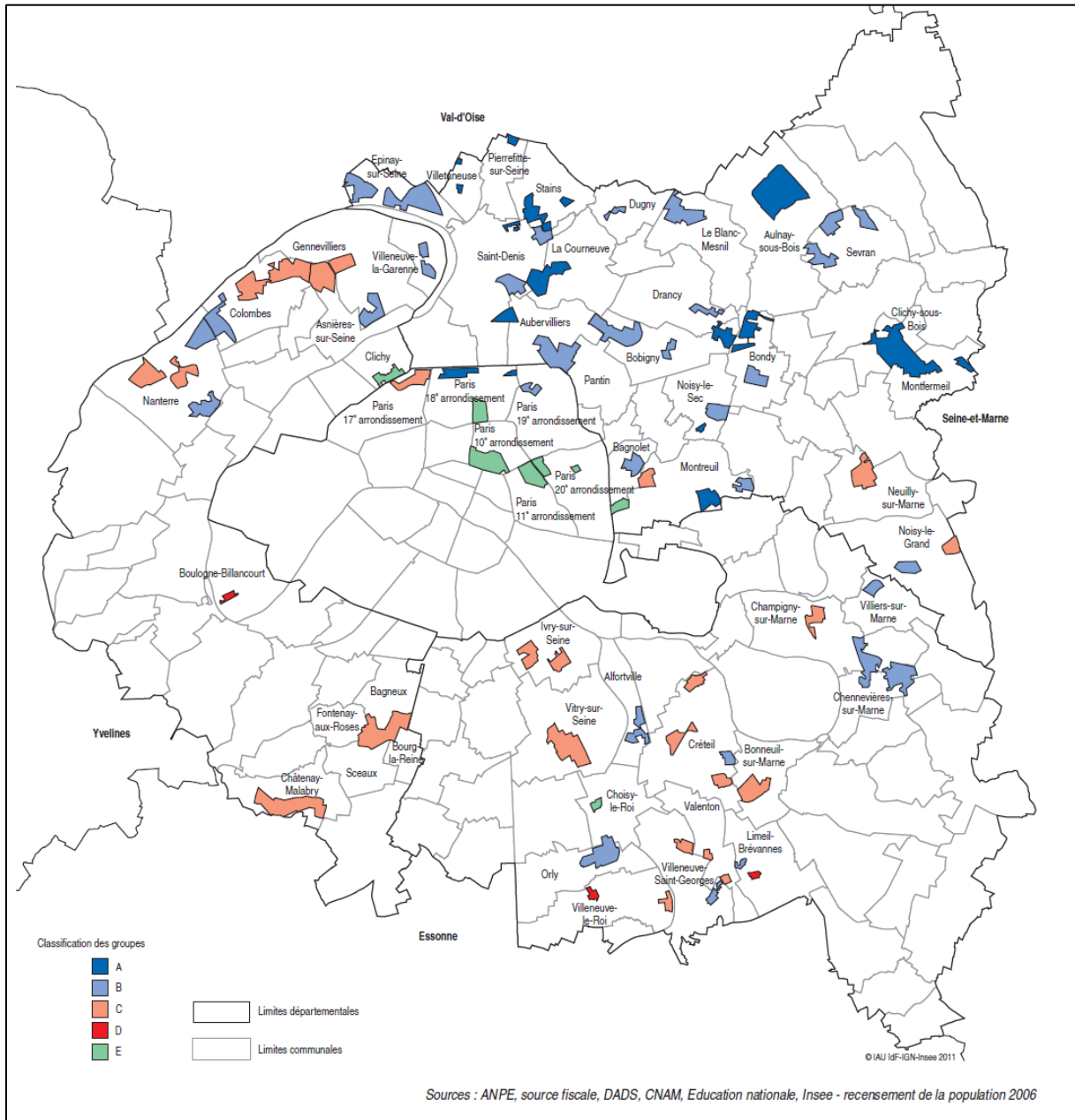
	Paris	Hauts-de-Seine	Seine-Saint-Denis	Val-de-Marne	Parisian Area
nb of ZUS area	9	16	36	23	84
Pop in ZUS areas	125 532	167 093	305 983	136 826	735 434
Pop in Departement	2 181 374	1 536 096	1 491 972	1 298 341	6 507 783
Pop ZUS/ Pop Dep. (%)	6%	11%	21%	11%	11%
nb of municipalities*	1 (20)	36	40	47	124 (143)
Area (km ²)	105	176	236	245	762
Density	20775	8728	6322	5299	8540
nb of IRIS	95	90	209	74	468
nb of IRIS ZUS	50	66	111	51	278
nb of IRIS un treated	45	24	98	23	190

Sources: Population Census (Insee 2006) and Musiedlak (Insee, 2011)

* For the City of Paris, the number in parenthesis is the number of districts

The Map 5-2 shows the location of the ZUSs in the Paris area according to five profiles defined by Musiedlak (2011) describing the variable degree of deprivation of the targeted zones. Group (A) is composed of ZUSs that are confronted with the most difficulties and they are located in the Saint-Denis department. Group (B) is composed of the ZUSs that are more distressed than the average ZUS in the Ile-de-France Region. Group (C) represents the intermediate rank i.e. the ZUSs that are less distressed than the average in the Ile-de-France Region. Group D is composed of the ZUSs that face fewer difficulties, and Group (E) is composed of IRISs considered as outliers and they are primarily located in the city of Paris.

Map 5-2: ZUS areas in Paris and its surrounding departements



At a finer scale, a ZUS covers several blocks where social interactions are engendered and grow. This means that in the same ZUS, different groups – i.e. living in different blocks – coexist. A ZUS is not necessarily socially homogenous but may be considered so when disaggregated data are unavailable at the block level. Many French data are provided by the

Insee at the IRIS scale,⁷⁸ which is a statistical unit aggregating some continuous blocks. The aggregation process must respect geographic and demographic criteria and have borders that are clearly identifiable and stable over the long term. Moreover, it respects the homogeneity of occupations: residential, economic or others. There are three types of IRIS unit: (i) the residential IRIS with a population generally between 1800 and 5000. The unit is homogeneous in terms of living environment and the boundaries of the unit are based on the major dividing lines provided by the urban fabric (main roads, railways, bodies of water,...); (ii) the business IRIS, with more than 1000 employees, with at least twice as many employees as other residents; (iii) the miscellaneous IRIS units: specific large zones that are sparsely inhabited and have a large surface area (leisure parks, ports, forests,...).⁷⁹ (Map 5-1 shows the Paris Area at the IRIS scale).

Because of these principles of homogeneity, the analysis of the ZUS program is more interesting at the IRIS than ZUS scale. Moreover, the sample becomes larger and more suitable for statistical and econometric analysis.

5.3.3 Identification of the sample of the IRIS units

Recall that the purpose of the treatment – the ZUS program – is to reduce the unemployment rate in ZUSs. Initially, evaluation of treatment requires two groups: the set of treated IRISs called the ZUS group and the set of untreated IRISs called the Non ZUS group that we take as the control group. To build the control group, we perform a matching with propensity score matching.⁸⁰ However, before that, we have to consider the type of statistical units available because in practice, a ZUS does not correspond exactly to a full set of IRIS units (Figure 5-1): on the fringes, only the part of the IRIS located within the ZUS is treated whereas the remaining part is untreated. To determine whether a particular IRIS should belong to the treated group or the control group, like Goffette-Nagot et al. (2012), we adopt the share of living space as a sorting criterion. We calculate the shares from the TOPO

⁷⁸ In fact, only towns with more than 10,000 inhabitants, and a large proportion of towns with between 5,000 and 10,000 inhabitants, are divided into several IRIS units. For the other towns, the IRIS unit corresponds to the town unit.

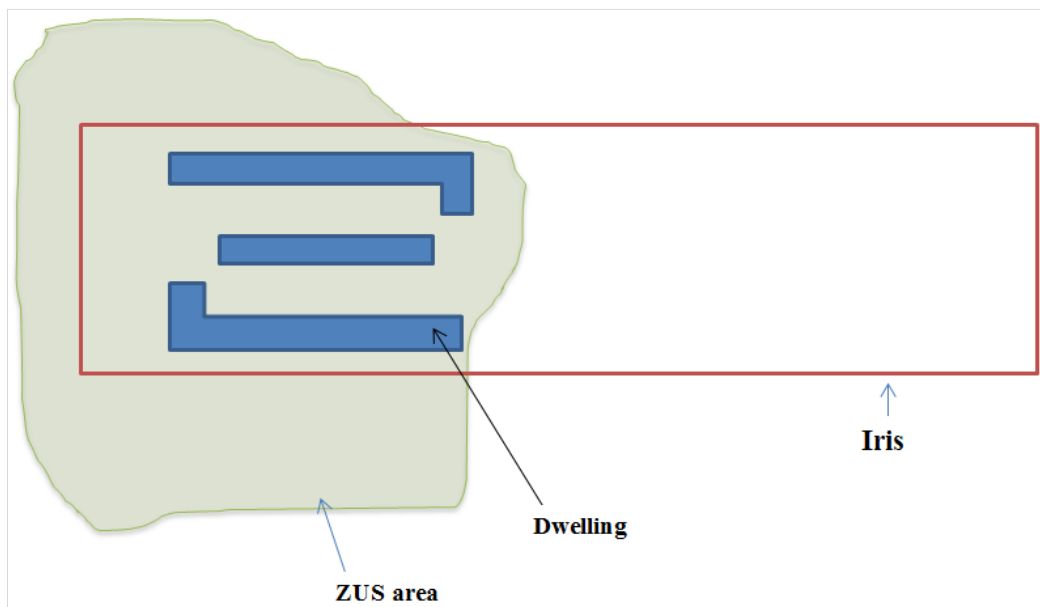
⁷⁹ The definition and details given in this chapter are drawn from the Insee web site.

<http://www.insee.fr/en/methodes/default.asp?page=definitions/iris.htm>

⁸⁰ We use R package MatchIt (Daniel E. Ho, Kosuke Imai, Gary King, Elizabeth A. Stuart (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, Vol. 42, No. 8, pp. 1-28. URL <http://www.jstatsoft.org/v42/i08/>)

database⁸¹ and any IRIS unit having a share of living space in the ZUS $\geq .50$ % is considered a ZUS IRIS and otherwise a Non ZUS IRIS. Let us note that considering this criterion (living area) rather than the proportion of the surface area of the IRIS in the ZUS is more suitable. The graph shows that if we had not considered the living area criterion, this IRIS would be excluded from the study because more than half of its surface area is outside the ZUS. In addition, with the surface area criterion, we could select all types of IRIS, Residential, Business, or Miscellaneous, although populated IRISs are of more interest for covering the ZUSs and evaluating the ZUS program. To implement our selection criterion, GIS was used.

Figure 5-1: ZUS and IRIS overlap



Source: Sileymane BA -2015

In the Paris area, the 87 ZUSs cover all or part of 450 IRIS units. After sorting, 278 residential IRISs are classified as ZUS units but three Business IRISs satisfying the sorting criterion are excluded.

To estimate the causal effect of the ZUS program, we need a control group composed of IRIS units located outside the ZUSs but with very similar characteristics to ZUSs. The matching is implemented with the propensity score matching.⁸² We estimate using a logit regression (the probability of one IRIS unit being classified in the ZUS group in accordance with certain

⁸¹ Database from the French National Geographical Institute

⁸² We use R package MatchIt (Daniel E. Ho, Kosuke Imai, Gary King, Elizabeth A. Stuart (2011). *MatchIt: Nonparametric Preprocessing for Parametric Causal Inference*. *Journal of Statistical Software*, Vol. 42, No. 8, pp. 1-28. URL <http://www.jstatsoft.org/v42/i08/>)

social and economic characteristics. The predicted probabilities are used to match the treated observations (ZUS) to the untreated observations (Non ZUS). The control group is composed of the untreated units that have very similar predicted probabilities to those of the treated units.

The Paris area is divided into 2557 IRIS units, 278 of which have been classified as ZUS IRISs. The explanatory variables X describing the social and economic status of IRISs are those traditionally used in empirical studies on urban segregation (Baumont, 2009; Goffette-Nagot et al., 2012). In the case of the ZUS program, these observable characteristics have to capture the deprived profile of IRIS units. Data are drawn from the Population Census database (Insee, 2006) and from the Fiscal database (2006). All variables are defined at the IRIS scale and concern the year 2006, i.e. ten years after the beginning of the ZUS program.

- Unemployment rate (UNEMPLOYMENT) is the percentage of unemployed persons in the working population of the IRIS.
- Proportion of young people (YOUNG) is defined as the share of people aged 15–24 years in the population aged 15–64 years.
- Proportion of no-diploma (NO- DIPLOMA) is defined as the share of people aged 15 years or more without any diploma in the the same age group.
- Proportion of low-rent housing (LRH) represents the share of social housing in the total housing in the IRIS.
- Proportion of foreign people (FOREIGNERS) represents the fraction of foreigners in the total population of the IRIS.
- Tax potential (TAX) represents the tax potential per inhabitant of the municipality weighted by the population of the IRIS.
- Median income (MEDIAN INCOME) is the annual median household income .

The estimated results of the logit model are presented in Table 5-4. As expected, the proportion of young people (YOUNG), the proportion of people without diplomas (NO DIPLOMA), the proportion of low-rent housing (LRH), the unemployment rate (UNEMPLOYMENT), and the proportion of foreign people (FOREIGNER) all have a positive and significant effect on the probability of the IRIS being a ZUS: IRIS units where such social indicators are high are more likely to be considered ZUS units. On the contrary, the economic characteristics – the tax potential (Tax) and the median income per consumer (MEDIAN

INCOME) – have the negative and significant expected effect on the probability of being considered a ZUS IRIS: IRISs with high indicators of wealth are less likely to be classified as ZUS units.

After matching, our sample includes 468 IRISs, with 278 IRISs considered as treated and classified in the ZUS group and 190 IRISs considered as not treated and classified in the Non ZUS control group (Map 5-3).

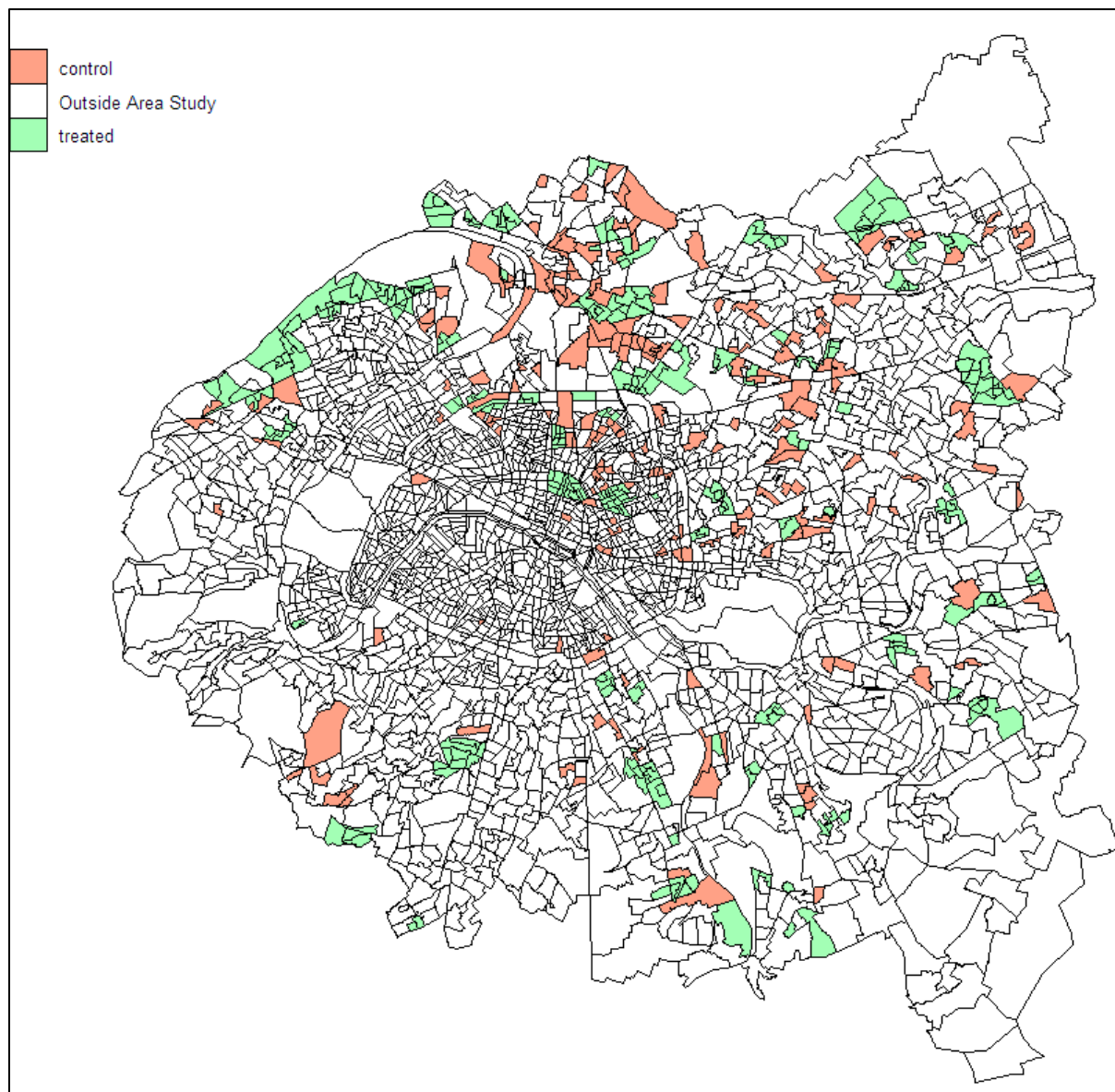
Table 5-4: Logit regression for Propensity Score Matching

Dependent variable: probability of IRIS being a ZUS unit

Variables (units)	Estimate (std.Error)
Intercept	17,468** (6,317)
YOUNG (%)	4,767** (1,564)
NO DIPLOMA (%)	2,688* (1,464)
LOW-RENT HOUSING (%)	1,541*** (0,397)
UNEMPLOYMENT RATE (%)	6,211** (2,309)
FOREIGNERS (%)	3,749* (1,674)
TAX PER CAPITA (€)	-0,006** (0,002)
MEDIAN INCOME (€)	-2,295*** (0,599)
N.obs	2557

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Map 5-3 : the spatial distribution of IRISs in ZUS and Non ZUS groups



Source : SIG de la politique de la ville 2006

Comparing the control group and the treated group, we can see that they have similar social attributes except for social housing Table 5-5. The unemployment rate is slightly higher in ZUSs than the sample average and amounts to 20 % on average in ZUSs against 18 % in Non ZUSs. The same ranges are observed for the proportion of the population without diplomas (NODIPLOMA): it is 35 % on average in the ZUS against 32 % in the Non ZUS group, for the proportion of young people (23 % on average in the ZUS group against 21 % for the control group), and for the proportion of foreign people (FOREIGNERS) (24 % on average in the ZUS group and 23 % in the Non ZUS group). On the contrary, the two groups

differ as regards social housing: the proportion of low-rent housing is very high in the treated group (57% on average) and lower in the control group (48 % on average). Finally, the economic attributes of Non ZUSs described by the tax potential (TAX) and the median household income per consumer (MEDIAN INCOME) are rather similar: the tax potential amounts to €39 in both groups and the ZUS units have a median income lower than the non ZUS units (€21,700 against €22,700).

Table 5-5: Descriptive statistics for observable characteristics

Variables (units)	Average in Non ZUS IRIS (st.dev.)	Average in ZUS IRIS (st.dev.)	Average for all IRIS (st.dev.)
YOUNG (%)	0,213 (0,040)	0,225 (0,048)	0,220 (0,045)
NO DIPLOMA (%)	0,317 (0,101)	0,351 (0,109)	0,337 (0,107)
LOW-RENT HOUSING (%)	0,481 (0,320)	0,570 (0,330)	0,534 (0,328)
UNEMPLOYMENT RATE (%)	0,180 (0,044)	0,196 (0,056)	0,190 (0,052)
FOREIGNERS (%)	0,229 (0,083)	0,244 (0,091)	0,238 (0,088)
TAX (€ PER CAPITA)	38,833 (37,515)	39,353 (35,204)	39,142 (36,121)
MEDIAN INCOME (€)	22746 (3900)	21691 (3825)	22119 (3886)
<i>Numbers of obs.</i>	<i>190</i>	<i>278</i>	<i>468</i>

Sources: Insee, Recensement de la population 2006 exploitation principale.

After having defined the control group and conditional variables, we present in the following section the estimated bounds of the treatment effect on the unemployment rate.

5.4 Estimation methods and results

The empirical application aims at estimating the effect of the ZUS program on unemployment: does the ZUS program successfully reduce unemployment in the IRISs classified as ZUSs. The variable of interest is the unemployment rate for the year 2006.

5.4.1 Estimation methods

In estimating the bounds of the effect of the ZUS program, we use the nonparametric regression method. This choice is driven by the characteristics of the data. All the covariates used are continuous variables. The distribution of the attributes varies across the IRIS units. Given the local characteristics of each IRIS, the relationship between the explanatory variables and the dependent variable is not necessarily either linear or homogeneous in all IRIS units.

The estimation by the nonparametric method therefore allows us to consider these aspects especially when we use very small statistical units and continuous variables implying a wider range of values than discrete variables.

In the standard – i.e. without spatial effect – nonparametric framework, the observed outcome is defined as:

$$y = m(x) + \epsilon \quad (5-20)$$

Where ϵ is the *i.i.d.* error term and $m(x)$ is a smooth function. $m(x)$ is not necessarily required to be linear since we are estimating the regression with nonparametric tools. In matrix form, the above equation can be written:

$$Y = M(X) + \epsilon \quad (5-21)$$

According to Gerolimetto and Magrini (2011), the nonparametric regression estimator for spatially dependent⁸³ redefines the outcome equation (5-5) as follows :

$$Y = M(X) + \epsilon \quad (5-22)$$

with $(X) = (I - \rho W)^{-1} X\beta$.

⁸³ For more details, see Gerolimetto and Magrini (2011).

“The function $M(X)$ consists of two components: one is $m(x)$, the function that characterizes the models in their original form, the second is a spatial factor that depends on the typology of spatial model that can possibly be the data generating process.” Gerolimetto and Magrini (2011, p. 25)

They stress that “estimating function M via SNP does not require assumptions on the functional form of M . This aspect is crucial since it means that no hypotheses are made either on m or the spatial factor that, multiplied by $m(X)$, generates $M(X)$ ”.

To estimate the bounds with the spatial effect, i.e. to control for spatial dependence, we use GAM models that can be used to treat spatial dependence in a nonparametric framework. GAM models allow for nonlinear relationships to be introduced between each covariate and the response variable.

Before including the spatial framework, we summarize the standard GAM model defined as follows (Wood, 2006):

$$\begin{aligned} y_i &= \beta_0 + \sum_{p=1}^P f_p(x_{ip}) + \varepsilon_i \\ &= \beta_0 + f_1(x_{i1}) + f_2(x_{i2}) + \dots + f_p(x_{ip}) + \varepsilon_i \end{aligned} \tag{5-23}$$

Where y_i is a response variable, x_{ip} a covariate, $f_p(\cdot)$ a smooth function, and the disturbances ε_i are i.i.d. $\mathcal{N}(0, \sigma^2)$.

To incorporate spatial dependencies, we can introduce two terms into equation (5-23). To capture the impact of the geographical location of the IRIS, we introduce the smooth term $f_2(long, lat)$, where $long, lat$ represent the coordinates of the IRIS unit. To capture the spillovers across IRIS units, we introduce an endogenous neighborhood effect (Manski, 1993) with the term WY . It enters endogenously because of the feedbacks between y and its spatial lag term (Basile, Benfratello, and Castellani, 2010).

Then the outcome observed is defined as follows:

$$y_i(x_i) = \beta_0 + f_1(x_i) + f_2(long, lag) + \rho \sum_{j \neq i} w_{ij} y_j + \varepsilon_i \quad (5-24)$$

In matrix form, equation (5-24) can be written:

$$Y(X) = \beta_0 + f_1(X) + f_2(long, lag) + \rho WY + \varepsilon_i \quad (5-25)$$

With conditional expectation:

$$E[Y|X = x] = \beta_0 + f_1(X = x) + f_2(long, lag) + \rho WY \quad (5-26)$$

Y is the vector of the outcome variable, X is a matrix of covariates, $long, lag$ represent the coordinates of the statistical unit, respectively the longitude and the latitude, and WY is the spatial autoregressive form of the dependent variable. WY is a vector given, for each outcome y_i of one IRIS i , its spatially weighted average defined on the neighborhood of the IRIS i . This variable allows spillovers from the ZUS program to be included and more specifically here it means that the unemployment rate in the IRIS i depends on the unemployment rate in its neighboring IRIS units.

The spatial weights matrix W is a (468x468) matrix in which the diagonal elements are (w_{ii}) equal to zero and the off-diagonal elements (w_{ij}) are defined as follows:

$$w_{ij} = \begin{cases} 1, & \text{if } dist_{ij} < 1000 \text{ m} \\ 0, & \text{otherwise} \end{cases}$$

Thus, each IRIS unit has as neighbors all the IRIS units located within a radius less than or equal to the threshold distance of 1000 m (distances are measured between the centroids of the IRIS units).

As indicated in Table 5-6, each IRIS has in average 7 neighbors. An IRIS can have both ZUS IRIS neighbors and Non ZUS IRIS neighbors. The neighborhood of a ZUS IRIS is composed by an average of 6 ZUS IRIS and 2 Non ZUS IRIS. The neighborhood of a Non ZUS IRIS is composed on average of 3 Non ZUS IRIS and 3 ZUS IRIS. However there are 28 IRIS that have no neighbors.

Table 5-6: Description of neighborhood matrices

Neighborhood' types	Average number of neighbors
ZUS IRIS- ZUS IRIS	6
ZUS IRIS- Non ZUS IRIS	2
Non ZUS IRIS- Non ZUS IRIS	3
Non ZUS IRIS- ZUS IRIS	3
All IRIS	7
Non neighbors	28

5.4.2 Results

We estimate the bounds of the treatment effect i.e. the effect of the ZUS program on the unemployment rate with three models: a-spatial, spatial-neighborhood (SNB), and Spatial Local and neighborhood (SLNB).

The spatial-neighborhood (SNB) specification includes neighborhood effects through the spatial autoregressive variable WY. It leads to spillover effects between all pairs of IRIS units. In the Spatial Local and Neighborhood (SLNB) specification, the local effect – linked to the geographical location of the IRIS unit within the Paris Area – is added. In each case, we estimate the bounds of the treatment effect conditionally for different covariates, taken individually: the proportion of young people (YOUNG), the proportion of people without diplomas (NO DIPLOMA), the proportion of foreigners (FOREIGNER), the proportion of social housing (LOW RENT HOUSING), the tax potential per capita (TAX), and the median income per consumer unit (MEDIAN INCOME).

5.4.2.1 Estimated bounds under the worst case

These first results are under the worst case i.e. when we consider what we can learn from the data alone. The estimated results are presented in Table 5-7. The first column concerns the estimation without spatial autocorrelation and spatial local effect. The second column shows the estimations with the neighborhood effect. In the third column, the bounds are estimated by taking into account both the neighborhood effect and the local effect.

First, we note that the estimated bound fails in all cases to provide any information about the treatment effects because the interval range includes the null. The treatment effect may be positive, negative, or zero. There is no indication whether the treatment effect is positive or negative. However, the estimation of the bounds with the neighborhood effect (SNB) increases the estimated bounds. The results show that, the lower bounds are shifted slightly toward zero and the upper bounds slightly away from zero.

In the third case, in which we consider both the neighborhood effect and the local effect, we do not observe any significant changes: one bound moves slightly while the other bound remains unchanged compared to the worst case.⁸⁴ The interval range always includes the null and the treatment effect remains undefined.

⁸⁴ Conditional on the proportion of young people, the lower bound increases by 1 %. For the proportion of foreigners, the upper bound increases by 1 %. For the proportion of social housing (Low-Rent Housing), the

5.4.2.2 Estimated bounds under the MTR assumption

In this section, we estimate the bounds under the MTR assumption. We suppose that the treatment – i.e. the ZUS program – improves the social and economic conditions of people living in the treated areas – the IRIS units of the ZUS group. More specifically, it is assumed that the policy leads to lower unemployment rates in ZUS areas. This is plausible because one objective of the program is to improve employment in such areas. Following Manski (1990) and Manski and Pepper (2000), we suppose that:

$$Y_1 \leq Y_0 \quad (5-27)$$

Where Y_1 the outcome (unemployment rate) is observed under treatment (ZUS group) and Y_0 is the outcome observed in the control group (Non ZUS group).

Combining equations (5–27) and equations (5–11) and (5–12) implies that:

$$E(Y_1|x, D = 1) \equiv \left[\begin{array}{l} \sup\{E(Y_1|x, D = 1)P(D = 1|x) + Y_1^l P(D = 0|x)\}; \\ \inf\{E(Y_1|x, D = 1)P(D = 1|x) + Y_1^u P(D = 0|x)\} \end{array} \right]$$

And

$$E(Y_0|X, D = 0) \equiv \left[\begin{array}{l} \sup\{E(Y_0|x, D = 0)P(D = 0|x) + Y_0^l P(D = 1|x)\}; \\ \inf\{E(Y_0|x, D = 0)P(D = 0|x) + Y_0^u P(D = 1|x)\} \end{array} \right]$$

Then the ATE under MTR is bounded by the lower bound B_{ATE}^{lMTR} and the upper bound B_{ATE}^{uMTR} :

upper bound declines by 1 %. For the tax potential, the lower bound declines by 1 % and for the median income both bounds remain unchanged.

$$\begin{aligned}
 B_{ATE}^{lMTR} &= \sup\{E(Y_0|x, D = 0)P(w = 0|x) + Y_0^l P(D = 1|x)\} \\
 &\quad - \inf\{E(Y_1|x, D = 1)P(D = 1|x) + Y_1^u P(D = 0|x)\}
 \end{aligned}
 \tag{5-28}$$

$$\begin{aligned}
 B_{ATE}^{uMTR} &= \inf\{E(Y_0|x, D = 0)P(D = 0|x) + Y_0^u P(D = 1|x)\} \\
 &\quad - \sup\{E(Y_1|x, D = 1)P(D = 1|x) + Y_1^l P(D = 0|x)\}
 \end{aligned}
 \tag{5-29}$$

As in the previous section, we have three cases: (i) the estimation under MTR without any spatial effect, (ii) the estimation under MTR with the neighborhood effect (SNB), and (iii) the estimation under MTR with both the neighborhood effect and the local effect (SLNB). All results are displayed in Table 5-8.

shows that under the MTR assumption, the estimated bounds are more informative conditional on three covariates: the treatment effects are negative – i.e. reduce the unemployment rate – conditional on the proportion of people without diplomas, the proportion of foreigners, and the median household income. Conditional on the other three covariates (the proportion of young people, the proportion of social housing, and the tax potential), the treatment effects remain undefined and we cannot appraise the impact of the ZUS program.

When we take the spatial effects into account, the results remain unchanged: the treatment effect is significant and negative (or not significant) once we have controlled for the same covariates. If the treatment effect is negative, the ZUS program successfully reduces the unemployment rate conditional on some covariates and this means that the ZUS program has a positive effect on employment in ZUSs.

Conditional on the proportion of people without diplomas, (NO DIPLOMA), the ZUS program effect on unemployment ranges from -0.089 to -0.044 in the a-spatial case, from -0.091 to -0.049 in the neighborhood effect case, and from -0.088 to -0.050 in the spatial trend and neighborhood effect case. Since the interval range decreases from the first case to the third one (from 0.044 to 0.039) the estimates become better and more precise. Taking into account both the spillovers and the local trend improves the estimation of the treatment effect.

Conditional on the proportion of foreigners, (FOREIGNER), the ZUS program effect on unemployment is in the range -0.088 to -0.026 in the first case, -0.103 to -0.032 in the second

case, and -0.096 to -0.029 in the third case. The first of the three cases displays the narrowest range (0.062) and so the best estimated bounds.

Conditional on the median income (MEDIAN INCOME), the ZUS program effect on unemployment is in the range -0.098 to -0.004 in the first case, -0.107 to 0-.004 in the second case, and -0.099 to -0.012 in the third case. The third case has the tightest range (0.087) corresponding to the best estimated bounds.

Conditional on the proportion of young people (YOUNG), the ZUS program effect on unemployment is uninformative in the first two cases but informative in the third case. The ZUS program effect is in the range -0.110 to -0.005.

Finally, taking into account the effects of neighborhood and local effects can better adjust the bounds and reduce the interval of the treatment effect.

Conclusion

The paper has been designed to take into account the effects of interactions between individuals in estimating the treatment effect in general and the bounds of treatment in particular. The main goal has been to improve the estimations of such bounds.

We have used three models to estimate the bounds in two cases. In the first run of estimation corresponding to the worst-case or no-assumption bounds, the results show that the introduction of spatial effects has a weak impact on the estimations even if it narrows the range of the bounds. In the second run of estimation, we suppose the Monotone Treatment Response assumptions on the observed outcome in order to ascertain informative intervals, tight bounds, and the sign of the treatment. The findings suggest that a more accurate estimation of bounds is obtained by considering both the neighbor effects and the local effect.

The results of this paper support the idea that the treatment applied to some individuals can produce spillovers that affect outcomes in their neighborhoods. Thus, the effect of the treatment is not confined to the direct effect on the treated alone but it incorporates the indirect effects generated by the treatment throughout the Paris area. In the bounds approach, this additional information can improve the estimation of the causal effect.

The approach developed in the paper contributes to Manski's method on treatment effects by including the effect of interactions among individuals regardless of their status (treated or untreated). The methods developed may be applied to other types of non experimental public policy or other variables of interest such as the school failure rate.

The most severe limitation on our empirical study is that we have not made any distinction among ZUSs. We have not considered the different categories of ZUS, i.e. ZFU and ZRU. We limited ourselves to the first definition that considers that both ZFUs and ZRUs are ZUSs. However, this problem is taken into account partly through the local effects used in the estimates.

Further research might explore the effect of heterogeneity in the neighborhoods because, as shown on the Map 5-3, a treated or an untreated individual has two types of neighbors and so the interactions differ depending on the neighbors.

Table 5-7: Estimated bounds of the ZUS program effect on the unemployment rate under the worst case

	a-spatial		SNB		SLNB	
	lower	upper	lower	upper	lower	upper
YOUNG	-0.113	0.190	-0.111	0.193	-0.114	0.190
NO DIPLOMA	-0.111	0.193	-0.110	0.194	-0.111	0.193
LOW-RENT HOUSING	-0.114	0.190	-0.110	0.193	-0.114	0.189
FOREIGNER	-0.114	0.189	-0.112	0.191	-0.114	0.190
TAX POTENTIAL	-0.117	0.187	-0.113	0.19	-0.116	0.187
MEDIAN INCOME	-0.112	0.192	-0.11	0.193	-0.112	0.192

Table 5-8: Estimated bounds of the ZUS program effect on the unemployment rate under the MTR assumption

	s-spatial			SNB			SLNB		
	lower	upper	width	lower	upper	width	lower	upper	width
YOUNG	-0.131	0.016	0.146	-0.116	0.002	0.118	-0.110	-0.005	0.105
NO DIPLOMA	-0.089	-0.046	0.044	-0.091	-0.049	0.042	-0.088	-0.050	0.039
LOW-RENT HOUSING	-0.088	-0.026	0.062	-0.103	-0.032	0.071	-0.096	-0.029	0.067
FOREIGNER	-0.145	0.073	0.218	-0.131	0.037	0.168	-0.100	0.019	0.119
TAX POTENTIAL	-0.158	0.098	0.256	-0.134	0.052	0.186	-0.127	0.044	0.170
MEDIAN INCOME	-0.098	-0.004	0.094	-0.107	-0.004	0.103	-0.099	-0.012	0.087

Conclusion générale

Dans cette thèse, nous nous sommes intéressés aux effets d'interaction dans l'évaluation des politiques publiques non expérimentales. Ces effets d'interaction constituent un enjeu majeur dans une évaluation, car ils peuvent influencer à la fois le processus de sélection des bénéficiaires et la performance de la politique. Alternativement, ignorer ces effets pourrait conduire à des estimations biaisées de l'effet de la politique. Nous proposons dans les différents chapitres de la thèse, l'application des méthodes de l'économétrie spatiale pour intégrer les effets d'interaction dans l'estimation de l'effet d'une politique.

Le premier chapitre est axé sur une méthodologie permettant d'intégrer les effets d'interaction dans les méthodes économétriques traditionnelles d'évaluation.

Le chapitre 2 porte exclusivement sur les effets d'interaction dans le processus de prise de décision ou de participation à une politique publique en général et dans le cadre du Grand Paris en particulier. Nous avons défini deux types de voisinage à savoir un voisinage institutionnel (appartenance à un même EPCI) et un voisinage géographique. Les résultats de l'étude empirique montrent qu'il existe des effets de voisinage, la décision d'une commune d'intégrer un contrat de développement territorial-CDT dépend alors des décisions prises dans son voisinage.

Dans le chapitre 3, nous nous sommes intéressés à l'effet *ex ante* d'une politique publique et nous proposons son estimation par la méthode des doubles différences en considérant les effets d'interaction. L'étude empirique porte sur les effets *ex ante* du projet Grand Paris sur les dépenses d'équipement par habitant. Nous estimons l'effet du Grand Paris sur trois périodes. La période 2001-2008 pour estimer l'effet pré-annonce du Grand Paris, la période 2008-2012 pour estimer l'effet post-annonce du projet et enfin la période globale 2001-2012 pour estimer l'effet *ex ante* du projet.

Les résultats montrent que les effets estimés du Grand Paris sur les dépenses d'équipement sont différents selon les périodes. L'effet pré-annonce du Grand Paris est non significatif, ce qui signifie qu'il n'y a pas eu de comportements d'anticipation des communes par rapport à leurs dépenses d'équipement. En revanche, sur la période *post*

annonce, les résultats indiquent que la consolidation du projet semble avoir un impact positif et significatif (au seuil de 10%) sur les dépenses des communes CDT par rapport aux communes hors CDT. Quant à la période globale, l'effet *ex ante* du projet est également positif et significatif (au seuil de 5 %) sur les dépenses d'équipement des communes. Cela veut dire que le projet Grand Paris semble avoir un impact positif sur les dépenses des communes CDT.

Concernant les effets des interactions, ils sont présents sur la période post-annonce et sur la période globale. Ils sont significatifs et négatifs ; ce qui veut dire que la commune diminue ses niveaux de dépenses lorsque les communes voisines augmentent les leurs. Ceci peut être interprété comme étant une complémentarité entre les communes.

Le chapitre 4 porte sur l'intégration des effets d'interaction sous forme de dépendance spatiale dans l'estimation de l'effet de traitement avec la méthode en deux étapes à la *Heckman*. Les estimations mettent en exergue trois résultats clés. Le premier permet de mettre en évidence l'existence d'effets d'interdépendances spatiales dans le processus de sélection des individus. Les estimateurs de l'équation de sélection ont été utilisés pour calculer l'équivalent spatial de l'Inverse du Ratio de Mills corrigé pour contrôler le biais de sélection. Le deuxième résultat met en lumière à la fois l'existence des effets de voisinage et l'hétérogénéité de ce voisinage. Il montre que l'effet de voisinage est différent selon le type de voisinage considéré, car un individu selon son statut a deux types de voisins : des voisins traités et des voisins non traités. Enfin le troisième résultat révèle que la prise en compte des effets d'interaction permet de réduire le biais de 7,5 points passant de 8,7 % sans les interactions à 1,2 % avec les interactions.

Le dernier chapitre porte sur les méthodes non-paramétriques des bornes de Manski (1990). C'est une méthode moins contraignante que les méthodes paramétriques, mais comme ces dernières, elle ne prend pas en compte les effets de voisinage dans l'estimation des bornes de l'effet du traitement. Nous avons proposé dans ce chapitre une extension de cette méthode en intégrant les effets d'interaction. L'étude empirique porte sur les effets de la politique des Zones Urbaines Sensibles (ZUS) sur le taux de chômage à l'échelle des IRIS de Paris et de la petite couronne. Elle

permet d'établir plusieurs éléments. Tout, d'abord, sans hypothèses sur les bornes (*worst case*), la prise en compte des effets de voisinage permet une amélioration des bornes mais ne donne aucune indication sur le signe du traitement. En revanche, en ajoutant des hypothèses supplémentaires sur la sélection (*Monotone Treatment Selection*) et les résultats (*Monotone Treatment Reponse*), il semble que l'intégration des effets de voisinage et des effets des caractéristiques locales permet de réduire l'intervalle de l'effet de traitement. Ces estimations fournissent également des indications sur le signe du traitement conditionnellement à certaines caractéristiques.

Les différentes études empiriques utilisées pour tester notre méthodologie mettent en exergue l'existence des effets d'interaction entre les individus. Elles montrent que la prise en compte de ces effets permet d'améliorer l'estimateur de l'effet de la politique évaluée. Ces résultats permettent de conforter l'hypothèse de recherche selon laquelle la prise en compte des effets d'interaction permet d'améliorer l'estimateur de l'effet de traitement.

Ce travail de thèse apporte une contribution à l'amélioration des méthodes d'estimation de l'effet de traitement. Les méthodes proposées peuvent faire l'objet d'application sur d'autres types de politiques publiques non expérimentales (éducation, santé, emploi,...) ou d'autres programmes de développement comme dans les pays en développement.

Cependant, les limites de la thèse résident sur les études empiriques, qui donnent certes des résultats probants, mais ne peuvent pas conduire à des conclusions généralisées. Ainsi, une question récurrente concerne le choix de la matrice de voisinage. Nous avons systématiquement choisi de spécifier plusieurs schémas d'interactions pour ne pas se limiter au cas géographique : choix d'une logique économique ou institutionnelle par exemple.

En termes de recherches futures, des approfondissements peuvent être faits sur des aspects statistiques inférentielles pour donner aux estimateurs une validité statistique et quantifier du point de vue statistique l'apport de l'introduction des effets d'interaction dans l'estimation de l'effet de traitement (réduction de biais, la

variance,...). Il serait également important d'approfondir la structure des voisinages en identifiant des individus « incontournables » (leaders) dans la population cible. Le développement de matrices de voisinage hiérarchisées par exemple permettrait de compléter les préconisations sur le rôle de certains acteurs dans la réussite de la politique publique.

Bibliographie

- Abreu M, De Groot, H.L.F, Florax, R.J.G.M (2005). Space and growth: a survey of empirical evidence and methods. *Région et Développement*. 21: 13-44.
- Acemoglu, D., Johnson, S., Robinson, J. a. (2005). Institutions as a fundamental cause of long-term growth. *Handbook of Economic Growth*, 1, 385-472.
- Affuso, E. (2010). Spatial Autoregressive Stochastic Frontier Analysis: An Application to an Impact Evaluation Study. Working Paper, Auburn University, available at *SSRN Electronic Journal*.
- Allers, M., Elhorst, J. P. (2005). Tax mimicking and yardstick competition among local governments in the Netherlands. *International tax and public finance*, 12(4), 1-27.
- Alpert, A. (2012). *The Anticipatory Effects of Medicare Part D on Drug Utilization*.
- An, W. (2011). Models and methods to identify peer effects. *The Sage Handbook of Social Network Analysis*. London: Sage, 515-532.
- Angrist, J. D., Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht.
- Anselin, L., Florax, R. 1995. Small sample properties of tests for spatial dependence in regression models, in Anselin, L.,Florax, R. (eds.), *New Directions in Spatial Econometrics*,Springer, Berlin.
- Arbia G (2006) *Spatial econometrics: Statistical foundations and applications to regional convergence*. Springer-Verlag, Berlin.
- Ba, S., Baumont, C. (2015). Hétérogénéité des interactions et évaluation des effets de traitement : une approche par les effets de dépendance spatiale. *Revue d'Économie Régionale Urbaine*, mai(1), 105-147.
- Baird S, Bohren J-A, McIntosh C, Özler B (2014) Designing experiments to measure spillover effects, PIER Working Paper Archive 14-006, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.
- Bandiera, O., Rasul, I. (2006). Social networks and technology adoption in the Zambezia province of Northern Mozambique. *The Economics Journal*, 116(1957), 869-902.
- Barnow, B. S., Cain, G. G. S., Arthur, G. (1980). Issues in the Analysis of Selectivity Bias, discussion Papers, Institute for Research on Poverty, University of Wisconsin-Madison
- Basile, R., Benfratello, L., Castellani, D. (2010). *Location determinants of foreign firms in the Enlarged Europe : evidence from negative binomial additive model*.

- Baumont, C. (2009). Spatial effects of urban public policies on housing values. *Papers in Regional Science*, 88: 301–326.
- Baumont, C., Guillain, R. (2013). Interactions, spillovers de connaissance et croissance des villes européennes : Quel est le rôle de la géographie, du climat institutionnel et des réseaux des firmes multinationales ? *Région et Développement*, 38, 161-207.
- Beaucire, F., Drevelle, M. (2013). « Grand Paris Express » : un projet au service de la réduction des inégalités d'accessibilité entre l'Ouest et l'Est de la région urbaine de Paris ? *Revue d'Economie Régionale et Urbaine*, 2013-03, 439-462.
- Becouze, E., Bouriat, P., Kupecek, J., Mesure, F., Moulines, S., Pillais, M. et Stecker, P. (2013). Les relations entre communes et communautés entre continuités et ruptures. *La Gazette des communes-des départements-des régions, Cahier détaché n° 2 - 43 / 2197*.
- Bergemann, A., Caliendo, M., Van Den Berg, G. J., Zimmermann, K. F. (2011). The threat effect of participation in active labor the threat effect of participation in active labor market programs on job search behavior of migrants in Germany. *International journal of manpower*, 32(7), 777-795.
- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society.*, 36(2), 192-236.
- Billé, A. G., Arbia, G. (2013). Spatial discrete choice and spatial limited dependent variable models: a review with an emphasis on the use in regional health economics. Working Paper University G. d'Annunzio.
- Blundell, R., Francesconi, M., Van der Klaauw, W. (2011). Anatomy of policy reform evaluation: Announcement and implementation effects. In *Technical Report, Working Paper*.
- Bramoullé Y., Habiba D., Fortin B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, Vol. 150, pp. 41-55.
- Briant, A., Lafourcade, M., Schmutz, B. (2014). Can Tax Breaks Beat Geography? Lessons from the French Enterprise Zone Experience. *American Economic Journal: Economic Policy*, 7(2): 88-124.
- Brock, W. A., Durlauf, S. N. (2001). Interactions-Based Models, in *Handbook of Econometrics*, Vol. 5, ed. by J. J. Heckman and E. Leamer. Amsterdam: North-Holland, 3297–3380.
- Brock, W. A., Durlauf, S. N. (2003). Multinomial Choice with Social Interactions, SSRI Working Paper No. 2003-1, University of Wisconsin
- Brock, W. A., Durlauf, S. N. (2007). Identification of binary choice models with social interactions. *Journal of Econometrics*, 140(1), 52-75.

- Brodaty, T., Crépon, B., Fougère, D. (2007). Les méthodes microéconométriques d'évaluation : développements récents et applications aux politiques actives de l'emploi. *Économie et Prévision*, n°177 2007, 26.
- Brueckner, J. K. (1998). Testing for strategic interaction among local governments: the case of growth controls. *Journal of Urban Economics*, 44(3), 438-467.
- Brueckner, J. K. (2003). Strategic interaction among governments. *International Regional Science Review*, 26(2), 175-188.
- Brueckner, J. K., Saavedra, L. A. (2001). Do local governments engage in strategic property-tax competition ? by do local governments engage in strategic property-tax competition ? *National Tax Journal*, 203-229.
- Bunel, M., Horty, Y. L., Petit, P. (2015). Mathieu Bunel, Yannick L'Horty, Pascale Petit. *Urban Studies*, SAGE Publications (UK and US), 1-31.
- Calvo-Armengol, A., Patacchini, E., Zenou, Y. (2009). Peer Effects and Social Networks in Education. *Review of Economic Studies*, 76(4), 1239-1267.
- Campos, C., Pinto, D. X. (2010). Semiparametric Estimation of Peer Effects in Classrooms : Evidence for Brazilian Schools in 2003. *mimeo*.
- Carré, D., Levratto, N. (2013). Les entreprises du secteur compétitif dans les territoires. Les déterminants de la croissance. 78p.
- Case, A. (1992). Neighborhood influence and technological change. *Regional Science and Urban Economics*, 22(3), 491-508.
- Case, A., Rosen, H. S., Hines, J. R. J. (1993). Budget spillovers and fiscal policy interdependence evidence from the states. *Journal of Public Economics*, 52, 285-307.
- Chagas, a. L. S., Toneto, R., Azzoni, C. R. (2012). A Spatial Propensity Score Matching Evaluation of the Social Impacts of Sugarcane Growing on Municipalities in Brazil. *International Regional Science Review*, 35(1), 48-69.
- Cont, R., Löwe, M. (2010). Social distance, heterogeneity and social interactions. *Journal of Mathematical Economics*, 46(4), 572-590.
- Crépon, B., Ferracci, M., Jolivet, G., Berg, G. Van den. (2010). Analyzing the anticipation of treatments using data on notification dates. IZA Discussion Paper 5265, Institute for the Study of Labor (IZA).
- Cressie, N. (1993). Statistics for spatial data. Statistics for spatial data. Willey, New York.
- Cressie, N., Chan, N. H. (1989). Spatial modeling of regional variables. *Journal of the American Statistical Association*, 84(406), 393-401.

- Daniel C, Garcia A, Roche-Bruyn F, Ruiz G, Verlhac E, Sardais C (2010) Rapport de la mission Evaluation des mesures en faveur des zones de revitalisation rurale (ZRR), novembre 2009, 124 pages.
- De Haan, M. (2014). The effect of additional funds for low-ability pupils: a non-parametric bounds analysis. *The Economic Journal*, (2014), 1-22.
- Dinardo, J., Lee, D. S. (2011). Program Evaluation and Research Designs. Handbook of labor economics, 4, 463-536.
- Dubé, J., Legros, D., Thériault, M., Des Rosiers, F. (2014). A spatial Difference-in-Differences estimator to evaluate the effect of change in public mass transit systems on house prices. *Transportation Research Part B: Methodological* 64(6): 24-40.
- Einio, E., Overman, H. (2011). The Effects of Spatially Targeted Enterprise Initiatives : Evidence from UK LEGI. LSE Spatial Economics Research Centre Working Paper.
- Elhorst, J. P. (2010). Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis*, 5(1), 9-28.
- Feder, G., Murgai, R., Quizon, J. (2004). Sending farmers back to school: the impact of farmer field schools in Indonesia. *Review of Agricultural Economics*, 26(1), 45-62.
- Feld, L. P., Josselin, J.-M., Rocaboy, Y. (2002). Le mimétisme fiscal : une application aux Régions françaises . *Economie prévision*, 5(156), 43-49.
- Ferracci, M. (2010). Interactions within Markets Treatment Evaluation in the Case of Interactions within Markets. IZA Discussion Papers, N°4700 2010
- Fleming, M. M. (2004). Techniques for estimating spatially dependent discrete choice models. In *Advances in spatial econometrics* (pp. 145-168). Springer Berlin Heidelberg.
- Flores-lagunes, A., Schnier, K. E. (2012). Estimation of sample selection models. *Journal of applied econometrics*, 204(July 2010), 173-204.
- Freret, S. (2008). Essais empiriques sur les interactions horizontales en termes de dépenses publiques. Thèse de doctorat en sciences économiques, Université de Rennes.
- Galster, G. (2007). Neighbourhood Social Mix as a Goal of Housing Policy: A Theoretical Analysis. *European Journal of Housing Policy*, 7(1), 19-43.
- Gerfin, M., Schellhorn, M. (2006). Nonparametric bounds on the effect of deductibles in health care insurance on doctor visits - Swiss evidence. *Health Economics*, 15(9), 1011-1020.
- Gerolimetto, M., Magrini, S. (2011). *A spatial nonparametric analysis of local multipliers*.

- Ghosh, S. (2011). Participation in school choice: a spatial probit analysis of neighborhood influence. *The Annals of Regional Science*, 50(1), 295-313.
- Gignoux, J. (2006). Évaluations *ex ante* et *ex post* d ' un programme d ' allocations scolaires conditionnées au Mexique. *Economie prévision*, 174(3), 59-85.
- Givord, P. (2010). Méthodes économétriques pour l 'évaluation de politiques publiques. *Document de travail INSEE*, 80 pages.
- Givord, P., Rathelot, R., Sillard, P. (2013). Place-based tax exemptions and displacement effects: An evaluation of the Zones Franches Urbaines program. *Regional Science and Urban Economics*, 43(1), 151-163.
- Goetzke, F. (2008). Network effects in public transit use : evidence from a spatially autoregressive mode choice model for New York. *Urban Studies Journal*, 45(2), 407-417.
- Goffette-Nagot, F., Charlot, S., Dujardin, C., Havet, N., Sidibé, M. (2012). *Accès à l'emploi dans les territoires de la politique de la ville : un appariement entre emplois et populations*. Rapport final. Paris : Dares, 119 p. (Aap "Evaluation des effets des politiques ciblées sur des territoires". Convention de Recherche DARES 6746. Partenariat CNRS 048206).
- Goldberger, A. S. (2008). Selection bias in evaluating treatment effects: Some formal illustrations. *Advances in Econometrics*, 21(1), 1-31.
- Graham, B. (2011). Econometric methods for the analysis of assignment problems in the presence of complementarity and social spillovers. *Handbook of social economics*, 1, 965-1052.
- Griffith, D. A. (2004). A spatial filtering specification for the autologistic model. *Environment and Planning A*, 36(10), 1791-1811.
- Gundersen, C., Kreider, B., Pepper, J. (2012). The impact of the National School Lunch Program on child health: A nonparametric bounds analysis. *Journal of Econometrics*, 166(1), 79-91.
- Guyon, X. (2007). Statistique spatiale. Dans *Statistique Appliquée pour le développement en Afrique* (88 p.).
- Haining, R. P. (2004). *Spatial data analysis: theory and practice*. Cambridge University Press.
- Halloran, M. E., Struchiner, C. J. (1991). Study designs for dependent happenings. *Epidemiology*, 5, 331-338.
- Hartmann, W. R., Manchanda, P., Nair, H., Bothner, M., Dodds, P., Godes, D., ... Tucker, C. (2008). Modeling social interactions: Identification, empirical methods and policy implications. *Marketing Letters*, 19(3-4), 287-304.

- He, F., Zhou, J., Zhu, H. (2003). Autologistic Regression Model for the Distribution of Vegetation. *Journal of Agricultural, Biological Environmental Statistics*, 8(2), 205-222.
- Heckman, J. J. (1979). sample selection bias as a specification error. *Econometrica*, 1, 153-161.
- Heckman, J. J. (2001). Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of political Economy*, 109(4), 673-748.
- Heckman, J., Ichimura, H., Smith, J., Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica*, 66(5), 1017-1098.
- Heckman, J. J., LaLonde, R. J., Smith, J. A. (1999). The economics and econometrics of active labor market programs. In O. A. And & D. Card (éd.), *Handbook of Labor Economics* (Vol. III, p. 1-73). Amsterdam: Elsevier.
- Heckman, J. J., Vytlačil, E. J. (2007). Econometric evaluation of social programs, part I: Causal models, structural models and econometric policy evaluation. *Handbook of econometrics*, 6, 4779-4874.
- Hilal, M., Barczak, A., Tourneux, F.-P., Schaeffer, Y., Houdart, M., Cremer-Schulte, D. (2011). Typologie des campagnes françaises et des espaces à enjeux spécifiques (littoral, montagne et DOM) synthèse. n°12, 80 p.
- Holland, P. W. (1986). Statistics and Causal inference. *Journal of the American Statistical Association*, 81(396), 945-960.
- Horowitz, J. L., Manski, C. F. (1999). Nonparametric analysis of randomized experiments with missing covariate and outcome data. *Journal of the American Statistical Association* 95(449), 77-84.
- Huang, J., Van Den Brink, H. M., Groot, W. (2012). Does education promote social capital? Evidence from IV analysis and nonparametric-bound analysis. *Empirical Economics*, 42(3), 1011-1034.
- Hudgens, M. G., Halloran, M. E. (2008). Toward Causal Inference With Interference. *Journal of the American Statistical Association*, 103(482), 832-842.
- Imbens, G., Wooldridge, J. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5-86.
- Jalan, J., Ravallion, M. (2003). Does piped water reduce diarrhoea for children in rural India. *Journal of Econometrics*, 112, 153-173.
- Jayet, H., Paty, S., Pentel, A. (2002). Existe-t-il des interactions fiscales stratégiques entre les collectivités locales. *Économie et Prévision* n°154, 154(3), 95-105.
- Jensen, R., Toma, E. F. (1991). Debt in a model of tax competition. *Regional Science and Urban Economics*, 21, 371-392.

- L'Horty, Y., Petit, P. (2011). Evaluation aléatoire et expérimentations sociales. *Revue française d'économie*, XXVI(1), 13-48.
- L'Horty Y., Sari F., (2013). Le Grand Paris de l'emploi : L'extension des infrastructures de transport peut-elle avoir des effets positifs sur le chômage local ? *Revue d'Economie Régionale et Urbaine*, 2013-03, 461-489.
- Lacombe, D. J. (2004). Does Econometric Methodology Matter? An analysis of public policy using spatial econometric techniques. *Geographical Analysis*, 36(2), 105–118.
- Le Gallo, J. (2002). Économétrie spatiale : l'autocorrélation spatiale dans les modèles de régression linéaire. *Economie et prévision*, 155, 139-157.
- Le Gallo, J., Mutl, J. (2014). Autocorrélation spatiale des erreurs et erreurs de mesure : quelles interactions ? *Région et Développement*, 40, 38-52.
- Lechner, M. (1999). Nonparametric bounds on employment and income effects of continuous vocational training in East Germany. *Econometric Journal*, 2, 1-28.
- Lechner, M. (2001). Identification and estimation of causal effects of multiple treatments under the conditional independence assumption (pp. 43-58). *Econometric Evaluation of Labour Market Policies*, 13, 43-58. Physica-Verlag HD.
- Lee, L., Liu, X., Lin, X. (2010). Specification and estimation of social interaction models with network structures. *Econometrics Journal*, 13(2), 145-176.
- Leenders, R. T. A. J. (2002). Modeling social influence through network autocorrelation: constructing the weight matrix. *Social Networks*, 24(1), 21-47.
- LeSage J-P, Pace R-K (2009) *Introduction to spatial econometrics*, New York: CRC Press.
- Leprince, M., Paty, S., Reulier, E. (2005). Choix d'imposition et interactions spatiales entre collectivités locales : un test sur les départements français. *Recherches Économiques de Louvain – Louvain Economic Review*, 71(1), 67-93.
- Liu, X., Lee, L. F. (2010). GMM estimation of social interaction models with centrality. *Journal of Econometrics*, 159(1), 99-115.
- Lofredi P. (2007). Évaluation économétrique des effets de traitement et programmes de développement à ciblage géographique. Le cas des interventions économiques en faveur du développement rural, Thèse de Doctorat en Sciences économiques de l'Université de Bourgogne, Dijon, UMR Cesaer (Inra-Enesad).
- Lorenceau, A. (2009). L'impact d'exonérations fiscales sur la création d'établissements et l'emploi en France rurale : une approche par discontinuité de la régression. *Economie et statistique*, 427(1), 27-62.
- Maddala, G. (1983). Limited-dependent and qualitative variables in economics. *Cambridge University Press*, pp. 257-91.

- Madiès, T. (1996). Concurrence fiscale et politique publique locale. *Annuaire des collectivités locales*, 16(1), 205-224.
- Madiès, T., Paty, S., Rocaboy, Y. (2005). Externalités fiscales horizontales et verticales. *Revue d'économie politique* 115(1): 17.
- Manski, C. F. (1990). Nonparametric bounds on treatment effects. *The American Economic Review*, 80(2), 319-323.
- Manski, C. F. (1993). Identification of social endogenous effects: the reflection problem. *The Review of Economics Studies*, 60(3), 531-542.
- Manski, C. F. (2013). Identification of treatment response with social interactions. *The Econometrics Journal*, 16, 1-23.
- Manski, C. F., Nagin, D. S. (1998). Bounding disagreements about treatment effects: a case study of sentencing and recidivism. *Sociological Methodology*, 28(1), 99-137.
- Manski, C. F., Pepper, J. V. (2000). Monotone instrumental variables: with an application to the returns to schooling. *Econometrica*, 68(4), 997-1010.
- Manski, C. F., Sandefur, G. D., McInahan, S., Powers, D. (1992). Alternative estimates of the effect of family structure during adolescence on high school graduation. *Journal of the American Statistical Association*, 87(417), 25-37.
- Manski, C.F. (2000). Economic Analysis of Social Interactions. *Journal of Economic Perspectives*, 14(3), 115-136.
- Mariotti, M., Meinecke, J. (2009). *Nonparametric bounds on returns to schooling in south africa.mimeo.*
- Mayer, T., Mayneris, F., Py, L. (2012). The Impact of urban enterprise zones on establishments' location decisions: evidence from french ZFUs. Discussion Papers Université Catholique de Louvain. Institut de Recherches Économiques et Sociales (IRES).
- Mcintosh, C. (2008). Estimating treatment effects from spatial policy experiments: An Application to Ugandan Microfinance. *The Review of Economics and Statistics*, 90(1), 15-28.
- McMillen, D. P. (1992). Probit with spatial autocorrelation. *Journal of Regional Science*, 32(3), 335-348.
- McMillen, D. P. (1995). Selection bias in spatial econometric models. *Journal of Regional Science*, 35(3), 417-436
- Moffitt, R. A. (2001). Policy interventions, low-level equilibria, and social interactions. *Social dynamics*, 4(45-82), 6-17.
- Musiedlak, Y. (2011). Les ZUS franciliennes: un paysage contrasté. *Alapage*, Insee Île-de-France, No.356.

- North, D. C. (1991). Institutions. *The Journal of Economic Perspectives* 5 (1), 97-112.
- Orfeuill J.P. (2014). Le métro « Grand Paris » : un choix contestable. *Futuribles*, 402, 37-54
- Overman, H. G. (2002). Neighbourhood effects in large and small neighbourhoods. *Urban Studies*, 39(1), 117-130.
- Pinkse, J., Slade, M. E. (1998). Contracting in space: an application of spatial statistics to discrete-choice models. *Journal of Econometrics*, 85(1), 125-154.
- Plaign, C. (2012). Neighborhood and Economic Spillovers: "Four essays on the role of culture, institutions and geography", PhD Thesis, Université Libre de Bruxelles
- Prager, J-C., (2014). Une nouvelle vision de la métropole parisienne. *Futuribles*, 401,53-67.
- Rathelot, R., Sillard, P. (2008). Zones Franches Urbaines : quels effets sur l ' emploi salarié et les créations d ' établissements? *Economie et Statistiques* 415-416 , 81-96.
- Ravallion, M. (2005). Evaluating anti-poverty programs. *World Bank Policy Research Working Paper Series, Vol.1-90*
- Redoano, M. (2003). *Fiscal Interactions Among European Countries*. Warwick Economic Research Papers, 680, 1-31.
- Rosenbaum, P. R. (2007). Interference between units in randomized experiments. *Journal of the American Statistical Association*, 102(477), 191-200.
- Rosenbaum, P. R., Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Rubin, D. B. (1974). Estimating Causal effects ot treatmentst in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688-701.
- Schaltegger, C. A., Küttel, D. (2002). Exit , voice , and mimicking behavior : evidence from Swiss. *Public Choice*, 133, 1-23.
- Seppo, K., Karikallio, H., Pirttilä, J. (2008). Anticipating tax changes : evidence from the Finnish Corporate Income Tax Reform of 2005. *Fiscal Studies*, 29(2), 167-196.
- Sherman, M., Apanasovich, T. V., Carroll, R. J. (2006). On estimation in binary autologistic spatial models. *Journal of Statistical Computation and Simulation*, 76(2), 167-179.
- Shipan, C. R., Volden, C. (2008). The mechanisms of policy diffusion. *American journal of political science*, 52(4), 840-857.
- Schone K. (2010), Stratégies d'influences et politiques de maîtrise de la croissance locale. Thèse de doctorat en sciences économiques, Université de Bourgogne.

- Schone K., Koch W., Baumont C., 2013, Modeling local growth control decisions in a multi-city case: Do spatial interactions and lobbying efforts matter?, *Public Choice*, 154(1-2), 95-117.
- Schone K. (2014), Quelles sont les origines des interactions stratégiques entre communes ? Le cas des politiques d'urbanisme. *Revue Canadienne de Sciences Régionales* 37 (1), 1-9.
- Schroeder, D. A. (2010) *Accounting and Causal Effects: Econometric Challenges*. Springer.
- Smirnov, O. A. (2010a). Modeling spatial discrete choice. *Regional Science and Urban Economics*, 40(5), 292-298.
- Smirnov, O. A. (2010b). Spatial econometrics approach to integration of behavioral biases in travel demand analysis. *Journal of the Transportation Research Board*, 2157, 1-10.
- Smirnov, O. A., Egan, K. J. (2012). Spatial random utility model with an application to recreation demand. *Economic Modelling*, 29(1), 72-78.
- Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate? *Journal of the American Statistical Association*, 101(476), 1398-1407.
- Soetevent, A. R. (2006). Empirics of the identification of social interactions; an evaluation of the approaches and their results. *Journal of Economic Surveys*, 20(2), 193-228.
- Solé-Ollé, A. (2006). Expenditure spillovers and fiscal interactions: Empirical evidence from local governments in Spain. *Journal of Urban Economics*, 59(1), 32-53.
- Tchetgen, E. J. T., VanderWeele, T. J. (2012). On causal inference in the presence of interference. *Statistical methods in medical research*, 21(1), 55-75.
- Todd, P. E. (2008). Evaluating Social Programs with Endogenous Program Placement and Selection of the Treated. Dans Paul Schultz and Strauss (Éd.), *Handbook of Development Economics*, 4 (John A. ed, pp. 3847-3894). (S.1.) : Elsevier.
- Todd, P. E., Wolpin, K. I. (2010). *Ex Ante* Evaluation of Social Programs. *Annales d'Economie et de Statistique*, 91, 259-286.
- Topa, G. (2001). Social interactions , local spillovers and unemployment. *Review of Economic Studies*, 68, 261-295.
- VanderWeele, T. J., Jiang, Z., Chiba, Y. (2014). Monotone Confounding, Monotone Treatment Selection and Monotone Treatment Response. *Journal of Causal Inference*, 2(1), 1-12. doi :10.1016/j.biotechadv.2011.08.021.Secreted
- Ward, P. S., Florax, R. J. G. M., Flores-lagunes, A. (2014). Climate change and agricultural productivity in Sub-Saharan Africa: a spatial sample selection model. *European Review of Agricultural Economics*, 41(2), 199-226.

- Wildasin, D. E. (1988). Nash equilibria in models of fiscal competition. *Journal of Public Economics*, 35, 229-240.
- Wood, S. (2006). Generalized additive models: an introduction with R. CRC press.
- Yamazaki, S., Resosudarmo, B. P. (2008). Does sending farmers back to school have an impact? Revisiting the Issue. *The Developing Economies*, 46(2), 135-150.