

Hétérogénéité des interactions et évaluation des effets de traitement : une approche par les effets de dépendance spatiale

Sileymane Ba, Catherine Baumont

▶ **To cite this version:**

Sileymane Ba, Catherine Baumont. Hétérogénéité des interactions et évaluation des effets de traitement : une approche par les effets de dépendance spatiale. Revue d'économie régionale et urbaine, 2015, Espace, statistique et économétrie : Progrès et défis en science régionale, 1-2, pp.105-148. 10.3917/reru.151.0105 . halshs-01224433

HAL Id: halshs-01224433
https://shs.hal.science/halshs-01224433

Submitted on 8 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CAIRN.INFO
CHERCHER, REPÉRER, AVANCER

HÉTÉROGÉNÉITÉ DES INTERACTIONS ET ÉVALUATION DES EFFETS DE TRAITEMENT : UNE APPROCHE PAR LES EFFETS DE DÉPENDANCE SPATIALE

Sileymane Ba, Catherine Baumont

Armand Colin | « Revue d'Économie Régionale & Urbaine »

2015/1 mai | pages 105 à 148

ISSN 0180-7307
ISBN 9782200929817

Article disponible en ligne à l'adresse :

http://www.cairn.info/revue-d-economie-regionale-et-urbaine-2015-1-page-105.htm

Pour citer cet article :

Sileymane Ba, Catherine Baumont, « Hétérogénéité des interactions et évaluation des effets de traitement : une approche par les effets de dépendance spatiale », *Revue d'Économie Régionale & Urbaine* 2015/1 (mai), p. 105-148.
DOI 10.3917/reru.151.0105

Distribution électronique Cairn.info pour Armand Colin.
© Armand Colin. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

Hétérogénéité des interactions et évaluation des effets de traitement : une approche par les effets de dépendance spatiale

Heterogeneity of interactions and evaluation of treatment effects: a spatial dependence approach

Sileymane BA

Laboratoire d'Économie de Dijon
UMR 6307 CNRS et Université de Bourgogne
sileymane.ba@u-bourgogne.fr

Catherine BAUMONT

Laboratoire d'Économie de Dijon
UMR 6307 CNRS et Université de Bourgogne
catherine.baumont@u-bourgogne.fr

Mots-clés : biais de sélection, effet de traitement, interactions spatiales, SUITVA, ZRR

Keywords : selection bias, spatial interactions, SUITVA, treatment effect, ZRR

Classification JEL : C13, C25, C31, C52

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

Résumé

L'évaluation des effets d'une politique publique peut être confrontée à l'existence d'interactions entre les individus. Dans ce cas, nous interrogeons la validité des hypothèses CIA *et* SUTVA *dans le cas du modèle d'évaluation causal en deux étapes. Nous utilisons un cadre d'analyse spatiale pour caractériser les interactions entre les individus. Nous modélisons l'hétérogénéité des voisinages selon que les individus et leurs voisins sont bénéficiaires ou non-bénéficiaires de la politique. Le modèle d'évaluation combine deux spécifications. Un modèle probit avec erreurs spatialement auto-corrélées permet de tenir compte des facteurs inobservables impactant potentiellement les décisions des individus à participer à la politique. Un modèle spatial autorégressif sur les résultats permet d'intégrer l'hétérogénéité des voisinages et l'éventuel biais de sélection spatial. Nous utilisons la politique de revitalisation des zones rurales (*ZRR*) comme application empirique. Avec les modèles sans interactions, les effets moyens de la politique sur la création d'établissements sont significatifs et positifs mais ils ne sont plus significatifs lorsque l'on intègre les interdépendances spatiales.*

Abstract

In the specific context of the two stage model of causal effect evaluation, this article deals with the analysis of the validity of the assumptions assuming no interdependency. More precisely we question the CIA *and the* SUTVA *assumptions. We use a spatial analytical framework to characterize the interactions between individuals. We consider the heterogeneity of neighborhoods according to individuals and their neighbors, whether or not they are involved in the public policy. The evaluation model combines two models. A probit model with a spatially autocorrelated errors makes possible to take account for unobservable factors potentially affecting individuals' decisions to take part in the policy. A spatial autoregressive model on the outcome allows us to integrate the heterogeneity of neighborhoods and the potential selection bias spatially adjusted. We use the* ZRR *(revitalization of rural areas) policy as an empirical application. We show that if the average treatment effects of the policy on the creation of business establishments are significant and positive with models without interactions, these effects are not significant anymore when we consider spatial dependencies.*

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

- 1 -
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 politiques[1]. Les méthodes économétriques d'évaluation des politiques publiques et, parmi elles, celles des effets de traitement (Rubin, 1974 ; Heckman et Vytlacil, 2007) 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 socioéconomiques 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 Independance 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 contraire à s'en détourner. Deuxièmement, l'hypothèse Sutva – *Stable Unit Treatment Value Assumption* – (Rubin, 1990) 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.[2]

Partant du constat que les interactions entre les individus existent et font partie intégrante des comportements et résultats observés, il est également difficilement

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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 (Calvó-Armengol *et al.*, 2009 ; Moffit, 2001 ; Topa, 2001). Par ailleurs, les interactions spatiales mettront 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 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 que, 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.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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, 1979 ; Heckman et *al.*, 1998). 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 des 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 et 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 résultat, 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'interactions 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 de l'hypothèse Sutva au cas des interactions intergroupes.

Le développement de ces propositions est organisé dans la suite de l'article de la manière suivante. Dans la deuxième partie, 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

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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 les extensions au modèle d'évaluation des politiques publiques dans un contexte d'interactions spatiales et sociales.

- 2 -
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, 1979 ; HECKMAN *et al.*, 1998) sont fixés dans un premier temps, avant de préciser le rôle et les effets de la prise en compte des interactions dans ce modèle.

2.1. Le cadre conceptuel du modèle d'évaluation

2.1.1. Le modèle d'évaluation en deux étapes

Nous considérons le cas des modèles à 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 :

$$\Delta_i = \gamma_{i1} - \gamma_{i0} \qquad (1)$$

γ_{i1} est la variable de résultat potentiel si l'individu i a reçu le traitement et γ_{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.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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 :

$$T^* = Z\alpha + u \tag{2}$$

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 :

$$T_{ik} = \begin{cases} 1 \; si \;\; T_{ik}^* > 0 \\ \\ 0 \; si \;\; T_{ik}^* \leqslant 0 \end{cases} \tag{3}$$

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 socioéconomiques 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 (BRODATY et *al.*, 2007). À 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 :

$$Y_k = X\beta_k + \varepsilon_k \text{ avec } k = \{0,1\} \tag{4}$$

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.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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} :

$$\Delta^{ATT}(X,Z) = E\left(Y_1 - Y_0 \mid T = 1, X, Z\right) \tag{5}$$

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} :

$$\Delta^{ATE}(X,Z) = E\left(Y_1 - Y_0 \mid X, Z\right) \tag{6}$$

2.1.2. Le contrôle des interdépendances

Généralement, on stipule l'absence d'interdépendances au sein du modèle à deux niveaux. Premièrement, l'hypothèse CIA[3] – *Conditional Independence Assumption* – (IMBENS et WOOLDRIDGE, 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 SUTVA[4] – *Stable Unit Treatment Value Assumption* – (RUBIN, 1990) postule que le traitement d'un individu affecte uniquement le résultat de cet individu et n'a aucun effet sur le 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.

2.2. 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.[5] 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

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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). 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 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 au fond 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, *etc.*, 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

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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.

2.3. 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 bénéfices 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 (Maddala, 1983 ; Brodaty et *al.*, 2007), 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.[6] 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

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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.[7]

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 Ctr (Manski, 2013) ou « *partial interference assumption* » (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

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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, Yamazaki et Resosudarmo (2008) 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.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

- 3 -
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*[8] 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). À 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.

3.1. 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.

Formellement, si l'on considère le processus de sélection précédemment décrit par les Équations 2 et 3, le modèle structurel de la variable latente T^* du *Probit spatial* s'écrit de la manière suivante :

$$T^* = Z\alpha + u \quad \text{avec} \quad u = \lambda W u + \upsilon \tag{7}$$

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

où Z désigne un ensemble de variables explicatives, où les erreurs suivent un processus autorégressif spatial et où υ est un terme d'erreurs homoscédastiques et indépendantes. La variable latente T^* est donnée par l'Équation 3.

Si $(\mathrm{I} - \lambda \mathrm{W})$ est non singulière, on peut récrire à nouveau l'Équation 7 sous la forme suivante :

$$T^* = Z\alpha + (I - \lambda W)^{-1}\upsilon \tag{8}$$

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).[9]

Si $\lambda = 0$, on retrouve un modèle Probit standard avec absence d'interactions entre les individus dans leurs choix. Si $\lambda \neq 0$, 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 homoscédastique et la fonction de vraisemblance s'écrit de la manière suivante (LeSage et Pace, 2009) :

$$\mathrm{L}\left(\mathrm{T}^*,\ \mathrm{W}|\lambda, \alpha, \sigma^2\right) = \frac{1}{2\pi\sigma^{2(\mathrm{n}/2)}} \left|\mathrm{I_n} - \lambda \mathrm{W}\right| \mathrm{e}^{-\frac{1}{2\sigma^2}\left(\upsilon' \upsilon\right)} \quad \text{avec } \upsilon = \left(\mathrm{I_n} - \lambda \mathrm{W}\right)\left(\mathrm{T}^* - Z\alpha\right) \tag{9}$$

Soit pour la fonction de log-vraisemblance :

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

$$lnL = ln\ |I_n - \lambda W| - \frac{n}{2} ln(2\pi) - \frac{n}{2} ln\left(\sigma^2\right) - \frac{(T^* - Z\alpha)^{'} \left[(I_n - \lambda W)^{'} (I_n - \lambda W)\right]^{-1} (T^* - Z\alpha)}{2} \quad (10)$$

On obtient alors des estimateurs des paramètres α (Équation 11) et σ^2 (Équation 12) :

$$\widehat{\alpha} = \left[Z^{'} (I_n - \lambda W)^{'} (I_n - \lambda W) Z\right]^{-1} Z^{'} (I_n - \lambda W)^{'} (I_n - \lambda W) T^* \quad (11)$$

$$\widehat{\sigma}^2 = \frac{[(T^* - Z\widehat{\alpha})]^{'} (I_n - \lambda W)^{'} (I_n - \lambda W) (T^* - Z\widehat{\alpha})}{n} \quad (12)$$

En remplaçant les paramètres estimés $\widehat{\alpha}$ et $\widehat{\sigma}^2$ dans l'Équation 10, on obtient la fonction de log–vraisemblance concentrée (Équation 13) :

$$lnL_c = C - \frac{n}{2} \ln \left(\frac{(T^* - Z\widehat{\alpha})^{'} \left[(I_n - \lambda W)^{'} (I_n - \lambda W)\right]^{-1} (T^* - Z\widehat{\alpha})}{n} \right) + \quad (13)$$

$$|I_n - \lambda W|$$

$$\text{avec } C = -\frac{n}{2} \ln (2\pi) - \frac{n}{2}$$

Plusieurs auteurs soulignent que l'estimation de l'Équation 13 par la méthode du maximum de vraisemblance traditionnel n'est pas efficace et peu convergente (FLEMING, 2004 ; BILLÉ et ARBIA, 2013). Plusieurs autres méthodes sont proposées : la méthode *Expectation-Maximization* (EM) utilisée par MCMILLEN (1995, 1992), 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, 2010). L'estimateur de PINKSE et SLADE (1998), utilisé par FLORES-LAGUNES et SCHNIER (2012), est une méthode d'estimation pour controler 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.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

3.2. 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 l'hypothèse SUTVA n'est pas respectée du fait d'interactions au sein des groupes et entre les groupes et nous postulons des interactions endogènes au sens de MANSKI (1993). 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 (*cf.* Équation 18).

Considérons d'une manière générale, l'équation de résultats suivante :

$$\mathrm{Y} = \rho_{\mathrm{k1}}\, \mathrm{W}_{\mathrm{k}}^{1}\, \mathrm{Y} + \rho_{\mathrm{k0}}\, \mathrm{W}_{\mathrm{k}}^{0}\, \mathrm{Y} + \mathrm{X}\beta_{\mathrm{k}} + \varepsilon_{\mathrm{k}} \qquad (14)$$

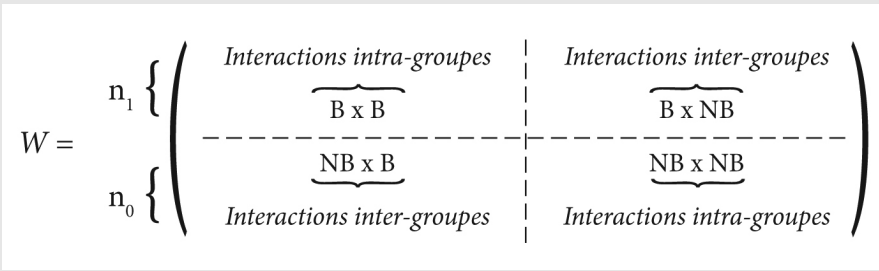
où Y représente la variable dépendante étudiée qui indique la performance de la politique (taux de réussite scolaire, taux d'insertion professionnelle, nombre de logements sociaux construits...). Si $k = 1$, on obtient l'équation de résultat pour les individus traités, et si $k = 0$, on retrouve l'équation de résultat pour les individus non traités (groupe de contrôle).[10]

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 s'ajouter aux interactions spatiales spécifiées par W compte tenu de l'ordonnancement 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 1) :

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

Figure 1 – Matrice de voisinage et typologie des interactions



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 :

$$W_1^1 = \begin{pmatrix} B \times B & 0 \\ 0 & 0 \end{pmatrix} \qquad W_1^0 = \begin{pmatrix} 0 & B \times NB \\ 0 & 0 \end{pmatrix}$$

$$W_0^1 = \begin{pmatrix} 0 & 0 \\ NB \times B & 0 \end{pmatrix} \qquad W_0^0 = \begin{pmatrix} 0 & 0 \\ 0 & NB \times NB \end{pmatrix}$$

Si la matrice $\left(\mathrm{I_n} - \rho_{k1}\,\mathrm{W_k^1} - \rho_{k0}\,\mathrm{W_k^0}\right)^{-1}$est inversible, alors on peut écrire les résultats (*cf.* Équation 14) sous la forme réduite suivante :

$$\mathrm{Y} = \left(\mathrm{I_n} - \rho_{k1}\,\mathrm{W_k^1} - \rho_{k0}\,\mathrm{W_k^0}\right)^{-1} \mathrm{X}\beta_k + \left(\mathrm{I_n} - \rho_{k1}\,\mathrm{W_k^1} - \rho_{k0}\,\mathrm{W_k^0}\right)^{-1} \varepsilon_k \qquad (15)$$

d'espérance μ_k et de matrice de variance-covariance Ω_k :

$$\mu_k = \left(I_n - \rho_{1k}\,W_k^1 - \rho_{0k}\,W_k^0\right)^{-1} \mathrm{X}\beta_k \qquad (16)$$

$$\Omega_k = \left(\mathrm{I_n} - \rho_{k1}\,\mathrm{W_k^1} - \rho_{k0}\,\mathrm{W_k^0}\right)^{-1} \left[\left(\mathrm{I_n} - \rho_{k1}\,\mathrm{W_k^1} - \rho_{k0}\,\mathrm{W_k^0}\right)^{-1}\right]' \sigma^2_{\varepsilon_k} \qquad (17)$$

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 bénéfices attendus. Une

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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 (*cf.* Équation 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 ROSEMBAUM (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 État (situation *within*), et d'autre part avec les *counties* les plus proches des États 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 ($\widehat{\rho}_{10} \neq 0$ et/ou $\widehat{\rho}_{11} \neq 0$) et les non-bénéficiaires ($\widehat{\rho}_{00} \neq 0$ et/ou $\widehat{\rho}_{01} \neq 0$). 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 $\left(\mathrm{I_n} - \rho_{\mathrm{k1}}\,\mathrm{W}_{\mathrm{k}}^{1} - \rho_{\mathrm{k0}}\,\mathrm{W}_{\mathrm{k}}^{0}\right)^{-1}$: toute variation d'une variable explicative X_p aura un effet différent de $\widehat{\beta}_k$ sur la variable dépendante Y.[11]

$$\frac{\partial \mathrm{Y_k}}{\partial \mathrm{x_p}} = \left(\mathrm{I_n} - \rho_{\mathrm{k1}}\,\mathrm{W}_{\mathrm{k}}^{1} - \rho_{\mathrm{k0}}\,\mathrm{W}_{\mathrm{k}}^{0}\right)^{-1} \beta_{\mathrm{k}} \neq \beta_{\mathrm{k}} \qquad (18)$$

L'estimation de l'Équation 15 par la méthode des moindres carrés produit des paramètres β_k biaisés et non convergents si les paramètres spatiaux ρ_{*k} sont significativement 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 14 s'écrit comme suit, pour $k = 1$ ou $k = 0$:

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

$$\mathrm{L}\left(Y_k,\ \mathrm{X}|\beta_\mathrm{k},\ \rho_{1\mathrm{k}},\rho_{0\mathrm{k}},\sigma^2\right)=\left(2\pi\sigma^2\right)^{-\mathrm{n}/2}\left|\mathrm{I_n}-\rho_{\mathrm{k}1}\,\mathrm{W}_\mathrm{k}^1-\rho_{\mathrm{k}0}\,\mathrm{W}_\mathrm{k}^0\right|\exp\left(-\frac{1}{2\sigma^2}\varepsilon_\mathrm{k}^{'}\varepsilon_\mathrm{k}\right) \quad (19)$$

avec $\varepsilon_\mathrm{k}=\left(\mathrm{I_n}-\rho_{\mathrm{k}1}\,\mathrm{W}_\mathrm{k}^1-\rho_{\mathrm{k}0}\,\mathrm{W}_\mathrm{k}^0\right)\mathrm{Y_k}-\mathrm{X}\beta_\mathrm{k}=Y-\mathrm{X}\beta_\mathrm{k}$

et $Y=\left(\mathrm{I_n}-\rho_{\mathrm{k}1}\,\mathrm{W}_\mathrm{k}^1-\rho_{\mathrm{k}1}\,\mathrm{W}_\mathrm{k}^0\right)\mathrm{Y_k}$

La fonction de log-vraisemblance est égale à :

$$\ln\mathrm{L}=\ln\left|\mathrm{I_n}-\rho_{\mathrm{k}1}\,\mathrm{W}_\mathrm{k}^1-\rho_{\mathrm{k}0}\,\mathrm{W}_\mathrm{k}^0\right|-\frac{\mathrm{n}}{2}\ln(2\pi)-\frac{\mathrm{n}}{2}\ln\left(\sigma^2\right)-\frac{\left(Y-\mathrm{X}\beta_\mathrm{k}\right)^{'}\left(Y-\mathrm{X}\beta_\mathrm{k}\right)}{2\,\sigma_\varepsilon^2} \quad (20)$$

On obtient alors des estimateurs des paramètres β_k et σ_ε^2 :

$$\beta_k=\left(\mathrm{X}^{'}\mathrm{X}\right)^{-1}\mathrm{X}^{'}Y \quad (21)$$

$$\widehat{\sigma}_\varepsilon^2=\frac{\left(Y-X\widehat{\beta}_k\right)^{'}\left(Y-X\widehat{\beta}_k\right)}{n} \quad (22)$$

3.3. 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 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 6).

Cependant l'effet du traitement sur les traités Δ^{ATT} n'est pas observable car le contrefactuel $E\left(Y_0 \mid T=1,X,Z\right)$ 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 $E\left(Y_0|\,\mathrm{T}=0\right)$. On définit alors un nouveau paramètre Δ qui est l'estimateur de l'effet moyen du traitement sur les traités Δ^{ATT}.

$$\Delta=E\left(Y_1|\,T\ =1\right)\ -E\left(Y_0|\,T\ =0\right) \quad (23)$$

La différence entre l'effet moyen du traitement sur les traités Δ^{ATT} et son estimateur Δ représente le biais.

$$\begin{aligned}\Delta&=E\left(Y_1|\,T\ =1\right)\ -E\left(Y_0|\,T\ =1\right)+E\left(Y_0|\,\mathrm{T}=1\right)-E\left(Y_0|\,\mathrm{T}=0\right)=\\ &\Delta^{ATT}+E\left(Y_0|\,\mathrm{T}=1\right)-E\left(Y_0|\,T\ =0\right)=\Delta^{ATT}+BS\end{aligned} \quad (24)$$

Le terme $\left[E\left(Y_0|\,T\ =1\right)-E\left(Y_0|\,T\ =0\right)\right]$ (Équation 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

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

de sélection est présent lorsque le 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.

$$\Delta=\Delta^{ATT}=\widehat{\Delta}^{ATE}=E(Y_1|\,T\,=1)\;-E(Y_0|\,T\,=0) \tag{25}$$

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 ($(E(Y_0|\,T\,=1)-E(Y_0|\,\mathrm{T}=0))\neq 0$). 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.

À partir des écritures de Schroeder (2010), les expressions du biais de sélection $E\;(\xi_1\mid T=1)$ et $E\;(\xi_0\mid T=0)$ sont données sous leur forme spatialement ajustée par les écritures suivantes :

$$E(Y_1|\,T\,=1)\;=\mu_1+E\;(\xi_1\mid T=1)\;=\mu_1+E\;(\xi_1\mid \nu>-Z\alpha) \tag{26}$$

$$E(Y_0|\,T\,=0)=\mu_0+E\;(\xi_0\mid T=0)\;=\mu_0+E\;(\xi_0\mid \nu<-Z\alpha) \tag{27}$$

avec : $\mu_1=E(Y_1|\,T=1)=\left(\mathrm{I_n}-\rho_{11}\,\mathrm{W}_1^1-\rho_{10}\,\mathrm{W}_1^0\right)^{-1}X\beta_1$

et $\mu_0=E(Y_0|\,T=0)=\left(\mathrm{I_n}-\rho_{00}\,\mathrm{W}_0^0-\rho_{01}\,\mathrm{W}_0^1\right)^{-1}X\beta_0$

$(\xi_1,\;\nu)$ et $(\xi_0,\;\nu)$ suivent une loi normale de moyenne 0 et de matrices de variance-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 :

$$E\;(\xi_1\mid T=1)\;=\;E\;(\xi_1\mid \nu>-Z\alpha)\;=\;\eta_1\sigma_1\frac{\phi(-Z\alpha)}{1-\Phi(-Z\alpha)}\;=\;\eta_1\sigma_1\frac{\phi(Z\alpha)}{\Phi(Z\alpha)} \tag{28}$$

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

$$E\ (\xi_0 \mid T=0) \ = \ E\ (\xi_1 \mid \nu < -Z\alpha) \ = \ \eta_0\sigma_0 \frac{-\phi(-Z\alpha)}{\Phi(-Z\alpha)} \ = \ \eta_0\sigma_0 \frac{-\phi(Z\alpha)}{1-\Phi(Z\alpha)} \tag{29}$$

Les Équations 26 et 27 deviennent :

$$E\,(Y_1|\,T\ =1) = \mu_1 + \eta_1\sigma_1 \frac{\phi(Z\alpha)}{\Phi(Z\alpha)} \tag{30}$$

$$E\,(Y_0|\,T\ =0) = \mu_0 + \eta_0\sigma_0 \frac{-\phi(Z\alpha)}{1-\Phi(Z\alpha)} \tag{31}$$

En présence de bais de sélection dans un cadre spatial, l'effet causal de la politique devient finalement (Équation 32) :

$$\widehat{\Delta}^{ATT} \ = \ E(Y_1 \mid T=1) \ - \ E(Y_0 \mid T=0) \ = \ [\mu_1 + E\ (\xi_1 \mid T=1)] \ - \ [\mu_0 + E\ (\xi_0 \mid T=0)]$$

$$\widehat{\Delta}^{ATT} = (\mu_1 - \mu_0) + \left[\eta_1\sigma_1 \frac{\phi(Z\alpha)}{\Phi(Z\alpha)} - \eta_0\sigma_0 \frac{-\phi(Z\alpha)}{1-\Phi(Z\alpha)}\right]$$

$$\widehat{\Delta}^{ATT} = (\mu_1 - \mu_0) + (\eta_1\sigma_1 - \eta_0\sigma_0)\,\frac{\phi(Z\alpha)}{\Phi(Z\alpha)}$$

$$\widehat{\Delta}^{ATT} = \Delta^{ATE} + (\eta_1\sigma_1 - \eta_0\sigma_0)\,\frac{\phi(Z\alpha)}{\Phi(Z\alpha)} \tag{32}$$

- 4 -
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 intéressant 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

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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.

4.1. Cadrage de l'analyse : la politique des Zrr et échantillon d'étude

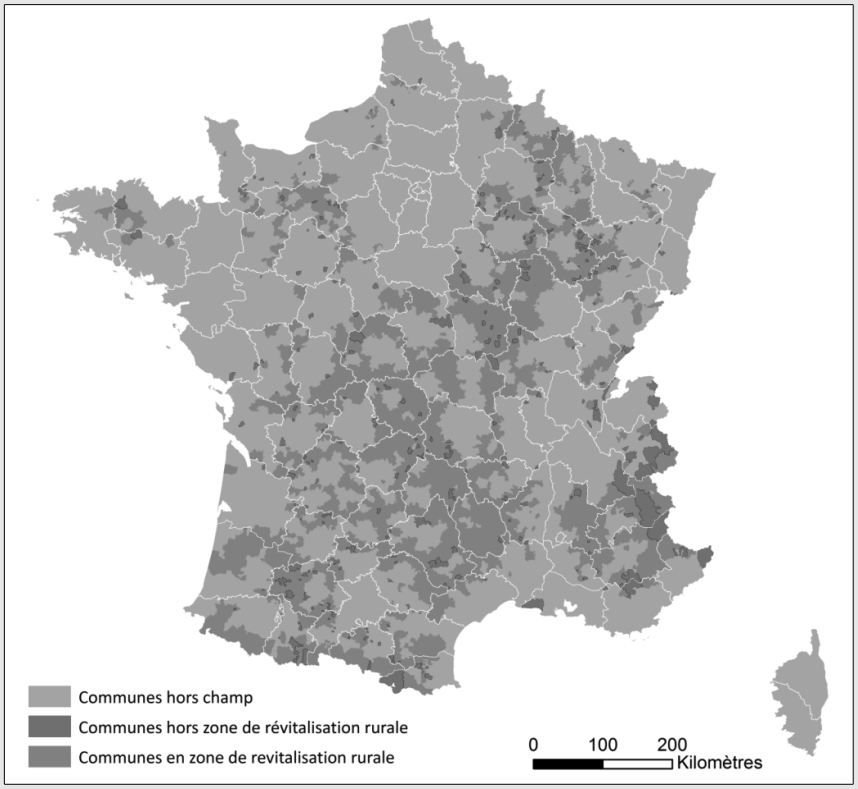
Les Zones de Revitalisation Rurale (Zrr), créées par la Loi d'Orientation pour l'Aménagement et le Développement du Territoire (Loadt) du 4 février 1995, désignent des territoires ruraux qui font face à des difficultés de développement. Cette loi vise à développer l'emploi et l'installation d'entreprises 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 socioéconomiques 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.[12]

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/km^2. 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.* Figure 2) : des communes classées Zrr sont entourées de communes

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

ZRR mais également de communes HZRR comme dans les zones littorales ou de montagne.

Figure 2 – 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/km^2.

Au total (*cf.* Tableau 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).

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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 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 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. À 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 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				
Industrie		299	2 755	3 054
Construction		528	3 148	3 676
Commerce, Transport, Services		2 537	13 352	15 889
Administration, santé, enseignement		644	1 967	2 611

Source : INSEE, RGP (2010) et Chiffres clés - Démographie des entreprises 2006-2011, Sirene. Calcul des auteurs.

4.2. 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.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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 2 fournit une présentation synthétique des variables utilisées dans l'équation de sélection, d'une part, et dans les équations de résultat, 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 afin de capter l'effet d'une polarisation éventuelle sur les communes de l'échantillon. La variable AIRU comprend plus précisément quatre modalités (*cf*. Tableau 3) permettant de repérer l'appartenance aux « *moyennes aires urbaines* », aux « *petites aires urbaines* »[13], 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*. Tableau 2), toutes proportions gardées du fait que nous nous situons dans des espaces ruraux et non urbains.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

Tableau 2 – Variables exogènes du modèle d'évaluation

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, *référence*)
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, *référence*)
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) : variable dummy Modalité de référence : Plaines et plateaux, prairies et forêts (Pp_Pf)	

**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.*

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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 socioprofessionnelles, permet d'apprécier un effet 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, nous avons pris en compte les catégories d'activités à l'échelle du bassin d'emplois.[14]

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 communes classées sont disponibles.[15] 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 dix 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 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.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

Tableau 3 – Typologies thématiques et répartition des communes

Typologie des aires urbaines*	Libellé	Hors ZRR	ZRR	Total
Communes isolées, hors influence des pôles	Isolée	429	4 852	5 281
Communes multi polarisées	MultiPol	258	2 231	2 489
Moyennes Aires urbaines	AUMoy	27	269	296
Petites Aires urbaines	AUPeti	34	267	301
Total		748	7 619	8 367
Typologie des paysages**	**Libellé**	**Hors ZRR**	**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	1 016	1 105
Plaines et plateaux, prairies et forêts	PP_PF	36	399	435
Collines et vallées, paysages diversifiés	CV_PD	78	986	1 064
Collines et vallées en prairies, peu artificialisées	CVP_PA	144	2 060	2204
Relief marqué, paysages diversifiés	RM_PD	91	1 221	1 312
Montagne, très forts reliefs et végétation semi-naturelle	MR_VSN	229	1 150	1 379
Total		748	7 619	8 367

** Source : INSEE, Nouveau Zonage en Aires Urbaines 2010, calculs des auteurs. ** Source : DATAR, Typologie thématique du champ « Paysages », 2011, calculs des auteurs.*

4.2.1. 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 3).

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

Figure 3 – Matrice de voisinage et typologie des interactions

$$W = \left(\begin{array}{c|c} \overbrace{\text{ZRR x ZRR}}^{\textit{Interactions intra-groupes}} & \overbrace{\text{ZRR x HZRR}}^{\textit{Interactions inter-groupes}} \\ \hline \underbrace{\text{HZRR x ZRR}}_{\textit{Interactions inter-groupes}} & \underbrace{\text{HZRR x HZRR}}_{\textit{Interactions intra-groupes}} \end{array} \right)$$

Le Tableau 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 (*cf.* Figure 2). 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 – Caractéristiques des matrices de voisinage

	Min	Max	Moyenne
W	21	950	479,16
Interactions intra-groupes			
Communes ZRR avec des voisins ZRR (ZRR - ZRR)	30	1 122	608
Communes HZZR avec des voisins HZZR (HZZR - HZZR)	37	1 045	482
Interactions inter-groupes			
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. Résultats des estimations[16]

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. À travers ces

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

estimations, nous donnons quelques interprétations des effets du programme ZRR sur le développement des communes.[17]

4.3.1. Équation de sélection

Le Tableau 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).[18]

Tableau 5 – Résultat de l'estimation de l'équation de sélection

	PROBIT standard	PROBIT spatial
	coefficient [1]	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	*ref*	*Ref*
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 %

*Variable dépendante : classement ZRR ou HZRR. Nombre d'observations : 8 367. Significativité : *** : 0,001 / ** : 0,01 / * : 0,05 / . : 0,1*

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 chances 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 chances 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

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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 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.

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.

4.3.2. Équations de résultats

Le Tableau 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).

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.[19]

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

Tableau 6 – Estimations des résultats pour les communes ZRR et hors ZRR

	Communes ZRR		Communes HZRR	
Variables	**Modèle a-spatial MCO [1]**	**Modèle SAR ML [2]**	**Modèle a-spatial MCO [3]**	**Modèle SAR 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*
Variables d'aménités : typologie des paysages[(1)]				
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***

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

Tableau 6 – (suite)

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

Variable dépendante : création d'établissements (en log).
Nombre d'observations : 8 367
(1) Pour la signification des variables d'aménités, se reporter au Tableau 3.
*Significativité : *** : 0,001 / ** : 0,01 / * : 0,05 / . : 0,1*

Nous analysons tout d'abord les résultats pour les communes bénéficiaires (*cf.* Tableau 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 voisines 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 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 inter-groupes. 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

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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 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 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'entrainement 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

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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.

4.3.3. 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 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 7 – Estimations des effets de traitements

	Modèle a-spatial		Modèle spatial	
	ATT	ATE	ATT	ATE
Nombre d'établissements estimés	2,675..	2,445..	3,7	3,6
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, toute chose égale par

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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.[20]

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).

- 5 -
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 d'invalidité de l'hypothèse SUTVA, 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.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

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 du traitement intègre la dimension spatio-temporelle des interactions sous la forme par exemple de matrices spatio-temporelles (Dubé et *al.*, 2013), cadre qui fera l'objet de recherches futures.

Remerciements

Cette étude s'intègre dans le programme de recherche « Approches économiques des dynamiques urbaines » du Puca (Plan Urbanisme Construction Architecture) et de la Maison des Sciences de l'Homme de Dijon, dirigé par Catherine Baumont. Les auteurs remercient Julie le Gallo, les participants du colloque 2013 de l'Afse (Aix-en-Provence) et les deux rapporteurs anonymes de la Reru pour leurs remarques et conseils sur les versions préliminaires de cet article.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

Annexes

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./km^2, ou bien (ii) dans un arrondissement ayant une densité inférieure ou égale à 33 hab./km^2, 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./km^2. Si ces densités de population sont inférieures à 5 hab./km^2, 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 socioéconomiques suivants.

- *Les critères socioéconomiques* 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 mars 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 quatre postes (industrie, construction, commerce-transports-service, administration publique-enseignement-santé-action sociale). Le champ d'activité agriculture est exclu de l'analyse.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

http://www.insee.fr/fr/themes/detail.asp?reg_id=99&ref_id=base-cc-demo-entreprises (Page consultée le 28 mars 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 un « 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 mars 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-th-matique-du-champ-paysages. (Page consultée le 24 mars 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 mars 2014)

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

Références bibliographiques

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.

Affuso E (2010) Spatial autoregressive stochastic frontier analysis: an application to an impact evaluation study. Working Paper, Auburn University, available at SSRN: http://dx.doi.org/10.2139/ssrn.1740382

Anselin L (1988) *Spatial econometrics: methods and models*. Kluwer Academic, Dordrecht.

Arbia G (2006) *Spatial econometrics: statistical foundations and applications to regional convergence*. Springer-Verlag, Berlin.

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.

Barnow B-S, Cain G-G-S, Arthur G (1980) Issues in the analysis of selectivity bias. In: Strosdorfer E, Farkas G (eds) *Evaluation Studies,* vol. 5, Sage Publications, Beverly Hills 42-59.

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.

Brock W-A, Durlauf S-N (2001) Interactions-based models. In: Heckman J-J, Leamer E-E (eds) *Handbook of econometrics*, edition 1, volume 5, Elsevier. 3297-3380.

Brodaty T, Crépon B, Fougère D (2007) Les méthodes micro-économétriques d'évaluation et leurs applications aux politiques actives de l'emploi. *Économie et Prévision*. 177: 93-118.

Calvó-Armengol A, Patacchini E, Zenou Y (2009) Peer effects and social networks in education. *Review of Economic Studies* 76(4): 1239-1267.

Case, A (1992) Neighborhood influence and technological change. *Regional Science and Urban Economics* 22(3): 491-508.

Daniel C, Garcia A, Roche-Bruyn F, Ruiz G, Verlhac E, Sardais C (2010) Rapport de la mission Évaluation des mesures en faveur des zones de revitalisation rurale (ZRR), novembre 2009, 124 pages.

Dinardo J, Lee D-S (2011) Program evaluation and research designs. In: Ashenfelter O, Card D (eds) *Handbook of Labor Economics*. vol. 4A, Elsevier Science B.V. 463–536.

Dubé J, Baumont C, Legros D (2013) Matrices de pondérations et contexte spatiotemporel en économétrie spatiale. *Canadian Journal of Regional Science/ Revue canadienne des sciences régionales.* 36(1/3): 57-75.

Einio E, Overman H (2011) The effects of spatially targeted enterprise initiatives: evidence from UK LEGI. LSE Spatial Economics Research Centre Working Paper.

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.

Fleming M-M (2004) Techniques for estimating spatially dependent discrete choice models. In: Anselin L, Florax R-J-G-M, Rey S-J (eds) *Advances in spatial econometrics. Methodology tools and applications.* Springer-Verlag, Berlin. 145-168.

Flores-lagunes A, Schnier K-E (2012) Estimation of sample selection models with spatial dependence. *Journal of Applied Econometrics* 20(4): 173-204.

Givord P (2010) Méthodes économétriques pour l'évaluation de politiques publiques, Document de Travail. G 2010/08. Direction des Études et Synthèses Économiques. Insee.

Goldberger A-S (2008) Selection bias in evaluating treatment effects: some formal illustrations. In: Fomby T, Carter H-R, Millimet D-L, Smith J-A, Vytlacil E-J (eds) *Modelling and evaluating treatment effects in econometrics. Advances in econometrics. Vol. 21.* Emerald Group Publishing Limited, 1-31.

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.

Heckman J-J (1979) Sample selection bias as a specification error. *Econometrica* 1: 153-161.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

Heckman J-J, Ichimura H, Smith J, Todd P (1998) Characterizing selection bias using experimental data. *Econometrica* 66(5): 1017-1098.

Heckman J-J, Vytlacil E-J (2007) Econometric evaluation of social programs, Part I: Causal models, structural models and econometric policy evaluation. In: Heckman J-J, Leamer E-E (eds) *Handbook of econometrics*, Volume 6B, Amsterdam and Oxford: Elsevier, North-Holland. Chapter 70. 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.* Datar Synthèse. n° 12-80.

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.

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. *Économie et Prévision* 155: 139-157.

Lechner M (2001) Identification and estimation of causal effects of multiple treatments under the conditional independence assumption. In: Lechner M, Pfeiffer F (eds) *Econometric Evaluation of Labour Market Policies*. Heidelberg: Physica. 43-58.

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-AgroSup Dijon).

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. *Économie et Statistique* 427-428 : 27-62.

LeSage J-P, Pace R-K (2009) *Introduction to spatial econometrics*, New York: CRC Press.

Maddala G-S (1983) *Limited-dependent and qualitative variables in economics*. Cambridge University Press, Cambridge.

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): S1–S23.

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.

Moffit R-A (2001) Policy interventions, low-level equilibria, and social interactions. In: Durlauf S-N, Peyton Y-H (eds) *Social Dynamics*. MIT press. 45-82.

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.

Rathelot R, Sillard P (2008) Zones franches urbaines : quels effets sur l'emploi salarié et les créations d'établissements ? *Économie et Statistique* 415-416 : 81-96.

Rosenbaum P-R (2007) Interference between units in randomized experiments. *Journal of the American Statistical Association* 102(477): 191-200.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

RUBIN D-B (1974) Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66(5): 688-701.

RUBIN D-B (1990) Comment: Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science* 5(4): 472-480.

SCHROEDER D A (2010) *Accounting and causal effects: econometric challenges.* Springer.

SMIRNOV O-A (2010) Modeling spatial discrete choice. *Regional Science and Urban Economics* 40(5): 292-298.

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.

TOPA G (2001) Social interactions, local spillovers and unemployment. *Review of Economic Studies* 68: 261-295.

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.

Notes

1 - Modernisation de l'Action Publique (MAP), décembre 2012.

2 - Les situations rencontrées dans un cadre expérimental sont nuancées par la suite dans la Section 2 de l'article.

3 - Ou encore *Unconfoundedness Assumption* pour IMBENS et WOOLDRIDGE (2009).

4 - On la retrouve également sous l'appellation *Individualistic Treatment Response* (ITR) pour MANSKI (2013).

5 - 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.

6 - 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.

7 - Le lecteur peut se reporter à l'Équation 8 pour visualiser cette propriété.

8 - On peut consulter FLEMING (2004) pour une présentation des différentes contributions sur ce sujet.

9 - 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.

10 - 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.

11 - 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.

12 - O, Historique des communes classées ZRR, Arrêté du 10 juillet 2013. COG0910 pour le classement du 31 décembre 2010.

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin

13 - 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).

14 - Cette 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 inexistants.

15 - 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.

16 - 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).

17 - 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.

18 - Les estimations des effets marginaux sont disponibles auprès des auteurs.

19 - 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.

20 - 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).

Document téléchargé depuis www.cairn.info - Université de Bourgogne - - 193.52.245.3 - 08/02/2017 10h04. © Armand Colin