

cee

CENTRE D'ÉTUDES DE L'EMPLOI

Juin
2010

Effets contextuels et effets de pairs :
quelles conséquences
sur la réussite scolaire ?

Sabina Issehnane,
Florent Sari

125

Document de travail

Effets contextuels et effets de pairs : quelles conséquences sur la réussite scolaire ?

SABINA ISSEHNANE

sabina.issehnane@cee-recherche.fr

Université Paris-XIII, CEPN et Centre d'études de l'emploi

FLORENT SARI

florent.sari@cee-recherche.fr

*Université Paris-Est Marne-la-Vallée, Centre d'études de l'emploi
et TEPP [FR CNRS n°3126]*

DOCUMENT DE TRAVAIL

N° 125

Juin 2010

ISSN 1776-3096
ISBN 978-2-11-098601-6

EFFETS CONTEXTUELS ET EFFETS DE PAIRS : QUELLES CONSÉQUENCES SUR LA RÉUSSITE SCOLAIRE ?

Sabina Issehnane, Florent Sari

RESUME

Bien que la question des effets de quartier ou de voisinage sur les comportements et les situations des jeunes ait fait l'objet de nombreuses études empiriques outre-Atlantique, tel n'est pas le cas en France. Ce travail essaye d'y apporter une contribution. Nous mobilisons les enquêtes *Emploi* pour étudier les comportements d'adolescents âgés de 18 ans et ce, sur une période allant de 1990 à 2002. Nous avons recours aux méthodes d'évaluation par appariements sélectifs afin de comparer les effets du voisinage sur les performances de deux groupes d'enfants équivalents, mais se distinguant par leur lieu de résidence. L'un vit dans un voisinage où le niveau socioculturel mesuré est faible et l'autre où le niveau socioculturel est important. Nous utilisons trois indicateurs différents pour mesurer la réussite scolaire : (1) abandonner ses études, (2) être diplômé et (3) être en retard scolaire. Nos résultats plaident majoritairement pour l'existence d'un effet négatif du contexte social sur la scolarité pour ceux qui vivent dans ces quartiers à plus faible capital socio-économique qu'ailleurs.

Mots-clefs : Effets de voisinage, réussite scolaire, méthodes par appariement sélectif.

JEL codes : I21, J24.

Contextual and Peer Effects : What Consequences on School Achievement?

Abstract

Although the issue of neighbourhood effects on the behaviour of young people has been the subject of many empirical studies in the United States, this is not the case in France. This work tries to fill this gap. We use the French employment surveys to study the behaviour of adolescents of 18 years old over the period 1990-2002. We develop matching methods to assess peer effects on the performance of two groups of children. One living in a neighbourhood where the socio-cultural level measured is low and the other where the socio-cultural level is important. We retain three different indicators to measure school achievement: giving up school (1) being a graduate (2) and educational backwardness (3). Our findings confirm a negative effect of social background on school achievement for those who live in neighbourhoods with lower socio-economic capital than others.

Key words: *Neighbourhood effects, school achievement, matching methods.*

INTRODUCTION

Les différences en termes de résultats scolaires sont souvent expliquées par le statut socioéconomique de la famille, le sexe et les capacités intellectuelles de l'élève. Pourtant, les stratégies résidentielles choisies ou subies et le rôle du contexte local contribuent vraisemblablement à renforcer le rôle des antécédents socio-économiques dans la réussite des enfants. Dans un article de 1987, Wilson suggère un processus par lequel le statut socio-économique du quartier affecte fortement les aspirations, attitudes et motivations de ceux qui évoluent dans ce quartier et, par conséquent, la réussite scolaire. En d'autres termes, tout se passe comme si les enfants étaient influencés par le comportement et le statut de ceux avec qui ils se socialisent ou de ceux qu'ils prennent comme modèle social.

Bien que la question des effets de quartier ou de voisinage sur les comportements et les situations des jeunes ait fait l'objet de nombreuses études empiriques outre-Atlantique (voir notamment les revues de littérature de Marpsat, 1999 ; Dietz, 2002), ce n'est pas le cas en France. En effet, encore peu de travaux sur données françaises cherchent à mettre en avant cet effet contextuel. L'une des rares études étant celle de Goux et Maurin (2005) qui utilisent judicieusement la structure aréolaire¹ des échantillons de l'enquête *Emploi* pour mettre en avant l'existence d'un effet causal significatif du contexte social. Une partie significative de l'échec scolaire semble ainsi s'expliquer par des effets de voisinage au-delà des caractéristiques socio-économiques individuelles.

Dans ce travail, nous mobilisons également la structure aréolaire de cette enquête et ceci afin d'estimer les effets causaux du quartier sur trois phénomènes distincts caractérisant la réussite scolaire : avoir abandonné ses études (1), avoir obtenu un certain diplôme (2) et être en retard scolaire (3). L'enquête *Emploi* nous permet d'étudier les comportements d'adolescents âgés de 18 ans et ce, sur une période allant de 1990 à 2002. Toutefois, mesurer l'effet du contexte local sur le développement des enfants soulève un problème de biais de sélection. Ici, c'est typiquement le cas dès lors que des individus ne se voient pas attribuer aléatoirement leur lieu de résidence. Pour corriger de ce biais nous avons recours aux méthodes d'évaluation par appariements sélectifs et effectuons des regroupements selon le score de propension (Rubin, 1974 ; Rubin et Rosenbaum, 1983). L'idée étant de comparer les effets du voisinage sur les performances de deux groupes de jeunes, toutes choses égales par ailleurs. L'un vivant dans un voisinage où le niveau socioculturel mesuré est faible et l'autre où le niveau socioculturel est plus important.

La méthode par appariement utilisée présente donc l'avantage de permettre une estimation de l'effet moyen d'un traitement donné sur une variable de résultat. Toutefois, si le biais de sélection est corrigé par la comparaison d'individus identiques sur leurs caractéristiques observables, il peut encore subsister un « biais caché ». En effet, s'il existe des variables inobservables qui affectent simultanément le fait d'être « traité » et la variable de résultat, un « biais caché » peut survenir qui remettrait en cause la robustesse des résultats obtenus avec la méthode par appariement (Rosenbaum, 2002). Ainsi, dans un deuxième temps, nous effectuons une analyse de sensibilité des résultats. Nous avons recours au test statistique de

¹ On entend par structure aréolaire le fait que les ménages interrogés dans l'enquête emploi soient tirés au sort par grappes de 20 ou 40 logements contigus (voir Goux et Maurin, 2005).

Mantel et Haenszel (1959) afin d'évaluer dans quelle mesure cette sélection sur « inobservables » biaise les résultats de nos estimations.

Ce travail est structuré de la manière suivante : la section 1 présente une brève revue de littérature sur les liens potentiels entre voisinage et réussite scolaire ainsi que sur les quelques récents travaux français ayant étudié la question. La section 2 expose la démarche empirique et présente quelques faits stylisés. La section 3 présente les résultats des estimations et apporte des éléments de conclusion.

1. EFFETS DE VOISINAGE ET RÉUSSITE SCOLAIRE : UNE REVUE DE LITTÉRATURE

Dans un premier temps, nous exposons brièvement les mécanismes potentiels qui lient le contexte local des individus à leur réussite scolaire au travers, notamment, de l'influence du statut des pairs ou de ceux qu'ils vont considérer comme modèle social, et ainsi de leur comportement. Dans un second temps, nous présentons quelques travaux français sur le sujet.

1.1. Effets de voisinage et réussite scolaire : quels liens potentiels ?

La littérature concernant les effets potentiels du voisinage sur le développement des enfants demeure assez récente. L'article de Wilson en 1987 s'impose en précurseur. Selon Wilson, les choix, les valeurs partagées dans les quartiers conduisent à une concentration de caractéristiques positives et/ou négatives. Lesquelles vont, en retour, affecter fortement les aspirations, attitudes et motivations de ceux qui évoluent dans ces quartiers. À partir de cette analyse, de nombreux travaux dans les sciences sociales ont essayé de mettre au jour les mécanismes et facteurs qui lient le contexte local et la trajectoire scolaire des enfants.

Pour un certain nombre d'auteurs, les enfants sont influencés par le comportement, le statut de ceux avec qui ils se socialisent ou qu'ils vont accepter comme modèles. On peut voir ici l'implication d'au moins trois groupes de théories différentes: les *théories de la contagion* sur l'influence des groupes de pairs comme mécanisme générateur des effets de voisinage, les *théories de la socialisation collective* sur le rôle central des modèles de rôle que les adultes de l'environnement proposent et les *théories de la compétition* fondées sur un concept de ressources limitées dans lequel les voisins et/ou les compagnons de classe se disputent des ressources rares (les bonnes notes, par exemple)².

Les *théories de la contagion* avancent l'idée selon laquelle, vivre dans un quartier dont le niveau socio-culturel mesuré est faible peut être un frein à la réussite scolaire et plus largement à l'acquisition de capital humain. La réussite d'un étudiant dépend fortement des caractéristiques socioéconomiques ou encore du comportement des autres étudiants qui l'entourent. Dans ces conditions, la concentration d'étudiants aux faibles capacités et en échec scolaire peut être susceptible d'agir négativement sur la période d'apprentissage. (Bénabou, 1993). De même, la *théorie épidémique* de Crane (1991)³ tend à montrer que la propension d'un adolescent à adopter un comportement est fortement corrélée à la proportion de personnes affichant déjà ce même comportement.

² Voir notamment la revue de littérature de Jencks et Mayer (1990) pour plus de détails sur les différents mécanismes et différentes théories à l'œuvre.

³ Cette théorie s'inscrit dans les théories de la contagion puisqu'elle en constitue une déclinaison ou un cas particulier.

Pour les *théories de la socialisation*, les adultes d'un voisinage ont une influence sur les enfants qui y résident et qu'ils côtoient. Les membres du voisinage vont exercer un contrôle social en aidant notamment les enfants à intérioriser les normes sociales. Le statut socio-économique du voisinage, leur capital culturel, ainsi que leurs aspirations, impactent sur le devenir des enfants. De plus, ils agissent comme des modèles, si bien que côtoyer des personnes ne symbolisant pas une certaine réussite scolaire ou sociale, ne motive pas ou peu un enfant à accroître ses propres chances de réussite durant son adolescence.

Les théories précédentes tendent à montrer les effets négatifs de l'appartenance à un voisinage aux caractéristiques négatives sur la réussite scolaire ; inversement elles tendent à montrer que les voisinages plus favorisés agissent positivement sur la réussite. Or, les *théories de la compétition* montrent, quant à elles, que la présence de voisins favorisés peut avoir des effets néfastes sur les autres membres du voisinage (Cutler and Glaeser, 1997). Concrètement, le fait de grandir dans un bon quartier peut être préjudiciable pour un enfant, en particulier dès lors que celui-ci est issu d'une famille plutôt défavorisée. L'hypothèse sous-jacente est que l'allocation de ressources rares est un jeu à somme nulle. En effet, si l'on admet l'idée que la réussite scolaire fait l'objet d'une compétition, les voisins aisés peuvent être considérés comme une menace accrue. Si les adolescents réagissent à l'échec en perdant de leur motivation, le fait d'être entouré de voisins aisés pourrait diminuer les chances de réussite.

Dans ce travail, nous tentons d'évaluer la pertinence des théories de la contagion et de la socialisation collective sur la réussite scolaire. Dans le sens où les théories de la compétition peuvent potentiellement biaiser les effets obtenus, nous essayons de contrôler de ce phénomène par la définition même que nous retenons d'un voisinage défavorisé. En effet, pour nous, un tel voisinage est un voisinage qui concentre de fortes proportions de populations fragiles (voir section 3). En ce sens la présence de familles aisées sera faible ou insignifiante ce qui tend à minimiser les problèmes de compétition tels qu'ils ont été décrits dans la littérature.

Enfin, il importe de distinguer différents effets de voisinage mis en avant par Manski (1993, 2000). Il distingue les effets de voisinage *endogènes*, les effets de voisinages *exogènes* et les « effets corrélés ». Les premiers correspondent à ces effets d'imitation et de renforcement qui sous-tendent les théories de la socialisation ou de la contagion. Les effets de voisinages exogènes (ou effets contextuels) surviennent lorsque le comportement d'un individu dépend des caractéristiques exogènes des individus du voisinage. C'est typiquement ces deux types d'effets qui nous intéressent dans ce travail. Les « effets corrélés », quant à eux, sont mis en avant dès lors que les individus d'un quartier tendent à avoir des caractéristiques similaires. Un phénomène qui peut s'expliquer par un processus de tri dans l'espace des individus. Or, ce tri spatial conduit également à un tri des comportements. Dans ce cas, on peut voir émerger des effets corrélés mais qui n'impliqueront pas nécessairement un lien de causalité entre le lieu de résidence et ce que l'on veut mesurer. Nous avons donc recours aux méthodes d'appariements sélectifs afin de contrôler d'une éventuelle auto-sélection des individus qui biaiserait la mise en avant d'un effet causal du contexte local sur la réussite scolaire.

1.2. L'influence des pairs sur la réussite scolaire en France

Si les travaux américains sur ce sujet abondent (voir notamment les revues de littérature de; Ginther, Haveman et Wolfe, 2000 ; Sampson, Morenoff et Gannon-Rowley, 2002 ; Dietz,

2002), les travaux français sur le sujet sont encore rares. On peut néanmoins citer le travail de Goux et Maurin (2005). Ces derniers utilisent la structure aréolaire des échantillons des enquêtes *Emploi* afin de définir de petits voisinages de quelques dizaines de logement. Ils utilisent deux expériences quasi-naturelles en comparant le sous-échantillon des jeunes qui viennent d'arriver dans une aire au sous-échantillon des adolescents qui y résident déjà, et en comparant ceux habitant en HLM depuis peu et ceux qui y habitent depuis un moment. À partir de simples estimations, ils montrent que : quelque soit le niveau socioculturel de leurs parents, les enfants sont davantage en retard à l'école lorsqu'ils habitent dans un voisinage où le taux de retard scolaire des autres enfants est élevé, ainsi que lorsque le niveau socioculturel des autres familles est faible.

Dans un autre registre, Piketty (2004) s'est intéressé à deux outils de politique éducative les plus fréquemment évoqués pour réduire les inégalités scolaires : la taille des classes et la ségrégation scolaire. Il montre un impact négatif de la ségrégation scolaire, mais néanmoins plus faible que celui des tailles de classe réduite. Un travail plus récent de Brodaty et Gurgand (2008) évalue les effets de pairs et l'influence du professeur sur la réussite scolaire dans les études supérieures. À partir d'une quasi-expérience dans les universités françaises, ils trouvent que ces deux effets sont significatifs. Plus précisément, les étudiants les plus faibles semblent tirer profit de la présence d'étudiants très performants dans le groupe. D'où l'intérêt de promouvoir la mixité dans les classes d'étudiants. Si ces travaux ne s'intéressent pas aux effets de voisinage à proprement parler, ils prennent en compte les effets de pairs qui relèvent d'une logique semblable.

Notre travail a donc pour objectif de montrer l'importance de considérer la dimension locale en matière de réussite scolaire. En effet, si les caractéristiques familiales gardent un rôle prépondérant, nous voulons néanmoins montrer qu'il serait erroné de ne pas tenir compte de ces effets contextuels.

2. LA DÉMARCHE EMPIRIQUE : DONNÉES, VARIABLES ET STRATÉGIE ÉCONOMÉTRIQUE

Dans ce travail, nous cherchons à estimer les effets causaux du quartier sur trois phénomènes distincts caractérisant la réussite scolaire : avoir abandonné ses études, avoir obtenu un diplôme et être en retard scolaire. L'enquête *Emploi* permet d'étudier les comportements d'adolescents âgés de 18 ans et ce, sur une période allant de 1990 à 2002. Nous avons recours aux méthodes d'évaluation par appariements sélectifs et effectuons des regroupements selon le score de propension⁴ afin de comparer les effets de contexte sur les performances de deux groupes d'enfants, toutes choses égales par ailleurs. L'un vivant dans un voisinage où le niveau socioculturel mesuré est faible et l'autre où le niveau socioculturel est plus élevé.

2.1. Les données et les principaux indicateurs

Afin de mesurer l'impact du voisinage sur la réussite scolaire, nous mobilisons donc les enquêtes *Emploi* de l'Insee pour l'ensemble de la période 1990-2002. Notre démarche est

⁴ Concrètement, un score de propension représente la probabilité, pour un individu à caractéristiques données, d'être exposé à un traitement. Dans notre cas le traitement correspond au fait d'habiter dans un quartier favorisé.

proche de celle de Goux et Maurin (2005) puisque nous utilisons la structure aréolaire de cette enquête⁵. Nous raisonnons effectivement au niveau des « aires » qui représentent des micro-quartiers abritant entre 20 à 40 logements. Dans ces « aires » toutes les personnes de plus de 15 ans sont enquêtées, c'est-à-dire environ 40 à 60 personnes. L'échantillon de départ comporte plus de 13 000 aires différentes, interrogées une fois, deux ou trois fois.

Sur l'ensemble de ces aires, nous nous intéressons aux ménages qui hébergent au moins un individu âgé de 17 ans et qui sont observés deux ans de suite. L'idée étant d'étudier l'impact du lieu de résidence pour des jeunes de 17 ans sur leur réussite scolaire un an plus tard (soit à l'âge de 18 ans). Ce choix garantit que nos individus se sont socialisés au moins une année avec les pairs qui l'entourent. Ce faisant, nous cherchons à estimer un effet « retardé » du lieu de résidence (au travers des effets de pairs potentiels) sur la réussite scolaire. Cette première contrainte réduit notre échantillon à près de 8 000 aires, pour un total de 17 203 ménages. Toutefois, les autres ménages du quartier sont conservés pour permettre la construction de nos indicateurs mesurant le niveau socioculturel du quartier.

Pour mesurer la réussite scolaire des adolescents à l'âge de 18 ans nous retenons trois variables différentes. Premièrement, nous regardons si l'individu a abandonné ou non ses études. C'est-à-dire, s'il est en formation au moment de l'enquête. Deuxièmement, nous regardons si l'individu a obtenu un diplôme. Plutôt que d'en retenir un en particulier, tel que le Brevet d'Etudes Professionnelles (BEP), le Baccalauréat (Bac) ou un diplôme équivalent à Baccalauréat + 2 ans dont la répartition en termes de réussite peut varier d'une génération à l'autre, nous retenons les trois. Nous créons ainsi une variable dichotomique qui informe si oui ou non l'individu a obtenu l'un de ces trois diplômes. Pour finir, nous cherchons à évaluer le retard scolaire. Pour cela, nous regardons quel niveau devrait raisonnablement atteindre un individu de 18 ans. Il ressort de quelques statistiques descriptives qu'un individu en retard est un individu qui n'a pas encore obtenu à cet âge au moins un diplôme tel que le Contrat d'Adaptation Professionnelle (CAP), un Brevet d'Etudes Professionnelles (BEP) ou encore le Baccalauréat et plus. Là encore, nous créons une variable dichotomique qui informe si oui ou non l'individu est considéré comme en retard.

Outre les caractéristiques individuelles des adolescents de notre échantillon nous ajoutons un certain nombre de variables permettant d'expliquer la réussite scolaire. D'une part, nous introduisons des caractéristiques sur le ménage dans lequel l'individu se situe : le nombre d'enfants, la catégorie socio-professionnelle du chef de ménage, son niveau de diplôme, sa nationalité (française ou non), le statut marital (en couple ou seul)... Nous cherchons aussi à prendre en compte les disparités territoriales, notamment entre les académies ou entre les villes. Nous introduisons donc des indicatrices de région ainsi que la tranche d'unité urbaine (commune rurale, urbaine de moins de 20 000 habitants, urbaine de 20 000 à 200 000 habitants...) dans laquelle se situe le micro-quartier.

De plus, il importe de contrôler du fait que notre échantillon d'étude – les jeunes de 17 ans – ne se focalise pas sur une seule génération. Certains de nos individus sont nés en 1974, d'autres en 1978 ou encore en 1983. Dès lors, si la réussite scolaire était plus moins forte d'une année sur l'autre, en raison de facteurs divers tels que l'implication des étudiants

⁵ Hormis le fait que le travail de Goux et Maurin et le notre s'intéressent aux effets de voisinage sur la réussite scolaire et qu'ils mobilisent la structure aréolaire de l'enquête emploi, les démarches sont très différenciées. Dans leur travail, Goux et Maurin mobilisent des régressions avec variables instrumentales pour analyser les effets de voisinage sur le redoublement avant l'âge de 11 ans. Ici, nous étudions les effets de voisinage sur différents indicateurs de réussite scolaire pour des adolescents de 18 ans, et ce en mobilisant les méthodes d'appariement sélectif (voir section suivante pour plus de détails sur cette méthode).

ou des professeurs sur une année donnée ou en raison de la démocratisation de l'accès au diplôme, les résultats obtenus peuvent être biaisés. Ainsi, nous devons contrôler du fait qu'obtenir un diplôme en 1973 n'est pas identique à l'obtention de ce même diplôme dix ans plus tard, indépendamment des caractéristiques des individus. L'introduction d'indicateurs temporels dans nos estimations permet de capturer l'hétérogénéité inobservée entre les années.

Enfin, nous cherchons à construire un indicateur qui permet de mesurer le niveau socio-culturel du quartier⁶. En d'autres termes, on cherche à identifier les quartiers perçus comme défavorisés et ceux qui ne le sont pas. Pour cela, on intègre différentes dimensions. Nous établissons dans chacun de ces micro-quartiers la part des actifs cadres ou professions intermédiaires, la part des personnes diplômées du supérieur, la part des actifs (non chômeurs), la part des familles vivant en couple et la part des ménages ne vivant pas dans un HLM⁷. Pour chacune de ces dimensions, nous attribuons une note variant de 1 à 10 sur la base d'un calcul de déciles pondérés⁸. En outre, ces diverses parts ont été calculées en ne prenant pas en compte la famille de l'adolescent considéré. Elles sont donc calculées à partir de l'ensemble des observations disponibles relatives à l'« aire » considérée.

Nous combinons ensuite les différentes notes obtenues par un simple calcul de moyenne. Dès lors, plus la note finale est élevée plus le quartier est perçu comme favorisé, et inversement. Nous excluons ensuite les « aires » ou quartiers appartenant aux déciles médians (du 4^e au 7^e décile). On obtient finalement, à l'une des extrémités de la distribution, les quartiers les plus défavorisés contre les plus favorisés, de l'autre. Par la suite, nous comparerons la réussite d'individus appartenant au premier groupe de quartiers avec celle d'autres individus appartenant au second groupe, toutes choses égales par ailleurs.

2.2. Quelques faits stylisés

Au total, notre analyse portera sur un échantillon de plus de 12 000 adolescents de 17 ans pour lesquels on mesure la réussite scolaire à l'âge de 18 ans. L'échantillon mélange différentes générations d'adolescents puisque ces derniers sont observés entre 1991 et 2002.

Les tableaux qui suivent proposent quelques statistiques descriptives de notre échantillon. En premier lieu, il importe de revenir sur notre définition d'un voisinage défavorisé ou, autrement dit, un voisinage où le niveau socioculturel mesuré est faible. Les caractéristiques qui ont été utilisées pour la mesure sont représentées dans le tableau 1. Ainsi les quartiers défavorisés sont ceux qui ont obtenu les plus mauvais classements pour les différentes caractéristiques retenues. On retrouve les plus faibles parts d'actifs cadres, de personnes diplômées du supérieur, de personnes ne vivant pas en HLM dans les voisinages ou quartiers définis comme défavorisés. D'autres indicateurs permettent moins de trancher entre les types de quartiers. C'est le cas de la part des personnes en activité ou de la part des ménages vivant

⁶ Il importe de préciser que nous mesurons ici davantage le capital social du quartier que son capital économique. Or ces deux notions sont assez différentes puisque il est vraisemblable que certains quartiers affichent un niveau socio-culturel important mais un niveau économique plus faible. Toutefois, l'introduction de caractéristiques telles que la part des cadres et professions intermédiaires dans notre indicateur constitue une bonne *proxy* de la richesse économique de l'aire.

⁷ Pour cette dernière variable, nous partons de l'hypothèse que les quartiers qui ont une proportion élevée de logements sociaux sont également des quartiers dont la proportion de ménages à faibles revenus ou à revenus modestes est plus importante.

⁸ Nous prenons en compte la taille des quartiers qui peut varier. On attribue donc une pondération à partir de la population totale du quartier.

en couple. Enfin, quand on s'intéresse à l'ensemble des aires, on retrouve des proportions intermédiaires puisque l'on inclut ici les quartiers favorisés, défavorisés et les quartiers aux classements médians.

Tableau 1 :
Caractéristiques du voisinage

	Total		Défavorisé		Favorisé	
	Moyenne	Écart-type	Moyenne	Écart-type	Moyenne	Écart-type
Part des actifs cadres ou prof. Interm.	7,5%	0,09	4%	0,04	14%	0,12
Part de la population en activité	58%	0,12	59%	0,11	57,5%	0,12
Part des personnes diplômées du supérieur	16%	0,14	10%	0,07	26,5%	0,17
Part des ménages vivant en couple	63%	0,11	64%	0,10	62%	0,13
Part des personnes ne vivant pas en HLM	83%	0,34	75%	0,39	94,5%	0,20
nombre d'aires	7992		2485		2489	

Source : Insee, Enquêtes « Emploi », 1990-2002.

Le tableau 2 permet de comparer les caractéristiques des individus qui habitent dans les différents types de quartier que l'on a précédemment identifiés. En premier lieu, on observe un certain nombre de caractéristiques pour lesquelles les populations des divers types de quartiers diffèrent peu. C'est le cas, notamment, pour les ménages qui vivent en couple, les ménages dont le chef est actif ou encore pour le sexe de ce dernier. Globalement, vivre dans un quartier défavorisé, moyen ou favorisé ne permet pas de discriminer les ménages sur ce type de caractéristiques. En revanche, dès lors que l'on s'intéresse au niveau de diplôme ou à la catégorie socio-professionnelle, le constat n'est plus le même. Bien entendu, de telles observations dépendent, en partie de la construction de notre indicateur d'aire « défavorisée ». Quoiqu'il en soit on retrouve les faits déjà évoqués dans le tableau 1 (c'est-à-dire une population plus fragile dans les aires « défavorisées ») : les personnes peu ou pas diplômées ou les catégories socio-professionnelles les plus basses sont plus fortement concentrées dans ces quartiers défavorisés. Ce résultat n'est pas forcément dû à la manière dont est construit notre typologie des « aires » puisque la grande majorité des diplômés et CSP exposées ici ne participent pas à cette typologie.

Par ailleurs, nous trouvons également que ce type d'aire abrite plus souvent des ménages avec des familles nombreuses (trois enfants ou plus). Nos variables qui mesurent la réussite scolaire sont également assez contrastées. Si la proportion d'adolescents ayant abandonné ses études est assez proche dans les différentes aires observées, ce n'est plus le cas lorsque l'on s'intéresse au retard scolaire et au fait d'avoir obtenu un diplôme. Globalement, les voisinages où le capital socio-culturel mesuré est faible semblent être les mêmes voisinages où la réussite scolaire est la plus faible.

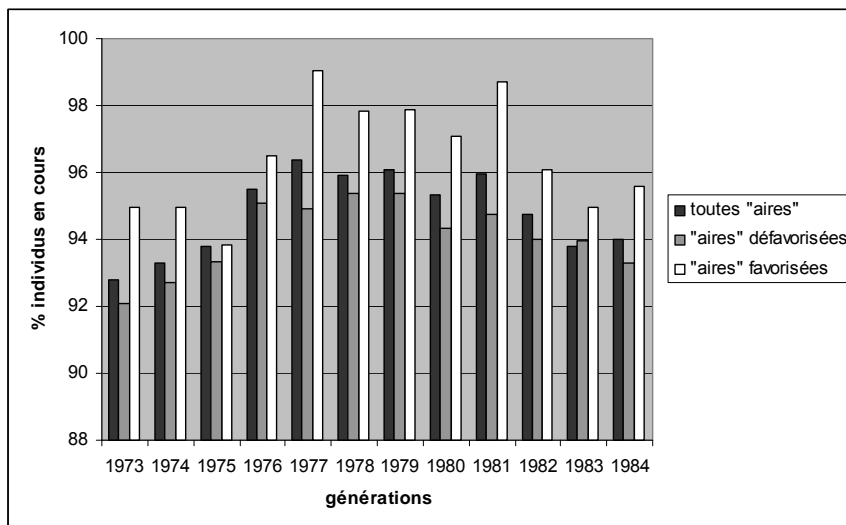
En outre, s'il existe un lien potentiel entre la localisation résidentielle et la réussite scolaire, il n'est pas certain que cela ait été toujours le cas pour les différentes générations de notre échantillon. Les graphiques 1, 2 et 3 comparent les différentes mesures de la réussite scolaire retenues pour les différentes générations et selon le lieu de résidence des parents. Ils permettent d'observer l'évolution des écarts au fil des années entre les individus vivant dans des quartiers favorisés et ceux vivant dans des quartiers évalués comme plus défavorisés.

Tableau 2 :
Caractéristiques individuelles selon le voisinage

	ensemble des "aires"		"aires" défavorisées		"aires" favorisées	
	Obs.	%	Obs.	%	Obs.	%
Caract. individuelles						
Homme	9476	51,3	4340	51,5	1965	51,2
Nationalité française	17504	94,8	7910	93,9	3698	96,3
réussite scolaire						
A abandonné les cours	17513	94,8	7918	93,9	3712	96,7
Est diplômé	12787	69,2	5516	65,5	2968	77,3
Est en retard	3128	16,9	1687	20	403	10,5
Caract. du ménage						
vie en couple	15285	82,8	7036	83,6	3112	81,1
Un enfant	3317	18	1344	15,9	822	21,4
Deux enfants	7319	39,6	3229	38,3	1642	42,8
Trois enfants ou plus	7600	41,2	3768	44,7	1313	34,2
Caract. du chef du ménage						
Homme	16195	87,8	7514	89,2	3267	85,4
En activité	15648	84,9	7134	84,7	3304	86,4
- Diplôme						
Diplôme supérieur	1733	9,4	384	4,6	808	21,1
Bac + 2	1148	6,2	404	4,8	367	9,6
Bac ou Brevet Pro.	1409	8,3	601	7,1	432	11,3
CAP, BEP	5644	30,6	2790	33,1	908	23,7
BEPC	1173	6,4	499	5,9	268	7,1
Aucun diplôme	7201	39,1	3743	44,5	1041	27,2
- Catégorie socio-professionnelle						
Chefs d'entr., cadres, prof. Int.	7377	40	2820	33,5	2162	56,5
Employés	2378	13	1051	12,5	475	12,4
Ouvriers	6443	35	3489	41,4	797	20,8
Autres	2268	12	1061	12,6	390	10,2
Caract. du lieu de résidence						
Commune rurale	5571	30,2	3178	37,7	504	13,2
Unité urbaine < 20 000 hab.	3274	17,8	2068	24,6	242	6,3
Unité urb. 20 000 à 200 000 hab.	3708	20,1	1828	21,7	645	16,9
Unité urbaine > 200 000 hab.	3675	19,9	837	9,9	1430	37,4
Agglomération parisienne	2238	12	510	6,1	1002	26,2
Total	18466		8427		3839	

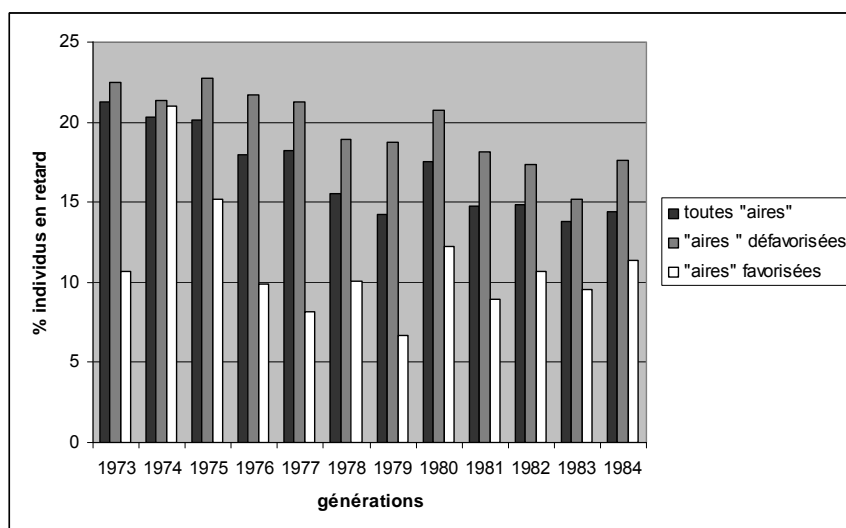
Source : Insee, Enquêtes « Emploi », 1990-2002.

**Graphique 1 :
Lieu de résidence et abandon scolaire**



Source : Insee, Enquêtes « Emploi », 1990-2002.

**Graphique 2 :
Lieu de résidence et retard scolaire**



Source : Insee, Enquêtes « Emploi », 1990-2002.

En premier lieu, nous mesurons les disparités en termes d'abandon scolaire entre les individus de 18 ans qui vivent dans les deux types de quartiers sur la période 1990-2002 (graphique 1)⁹. Globalement, il ressort de ce premier graphique que l'abandon scolaire a quelque peu diminué sur la période considérée. En effet, la part des individus encore en cours s'est accrue d'environ un point de pourcentage entre les individus nés en 1973 et ceux nés en 1984. Cependant, cette part a eu tendance à décroître entre 1977 et 1983, puisque l'on est

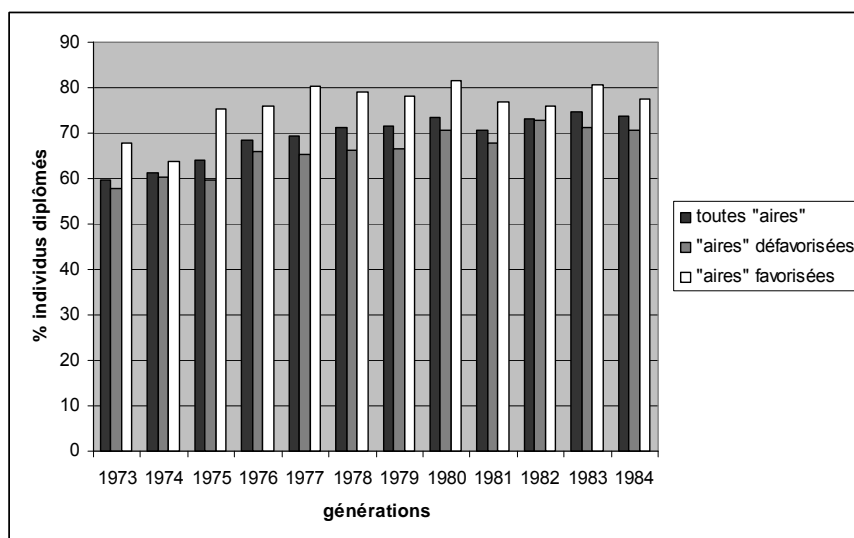
⁹ Ce qui correspond à des individus nés entre l'année 1973 et l'année 1984.

passé de plus de 96 % d'individus toujours en formation au moment de l'enquête à moins de 94 %. Il importe de préciser que les écarts observés ne sont pas forcément la résultante de disparités en termes de taille entre les différentes cohortes, puisque chacune des générations couvre approximativement 8 % de l'échantillon global.

Enfin, les écarts sont plus importants entre les individus vivant dans les aires défavorisées et ceux vivant dans les aires plus favorisées et ne semblent pas s'atténuer au fil des générations. En outre, la cohorte d'individus nés entre 1977 et 1981 apparaît comme la plus exposée aux inégalités en termes d'abandon scolaire.

Concernant le retard scolaire (graphique 2), il apparaît comme ayant globalement diminué entre la génération des individus nés en 1973 et ceux nés en 1984. La part des individus considérés comme étant en retard scolaire passe de plus de 20 % à moins de 15 % sur l'ensemble de la période. Pour les individus vivant dans les quartiers plus défavorisés, cette proportion est aussi en diminution, mais le phénomène demeure moins marqué et le retard scolaire reste encore largement plus élevé qu'il ne l'est pour la population totale (le retard scolaire dans ces quartiers oscille entre 22 % et 17 % sur la période). Pour les individus vivant dans des quartiers favorisés, hormis un choc pour la génération d'individus nés entre 1974 et 1975, on observe une faible proportion d'individus en retard scolaire puisque celle-ci demeure concentrée autour de 10 %. C'est-à-dire une proportion quasiment deux fois moins élevée que pour celle des individus vivant dans les quartiers défavorisés. D'une manière générale, ces quelques chiffres attestent d'une baisse tendancielle du recours au redoublement depuis 30 ans, avec notamment le développement de politiques de limitation du redoublement.

Graphique 3 :
Lieu de résidence et réussite scolaire



Source : Insee, Enquêtes « Emploi », 1990-2002.

L'évolution du système scolaire au fil des générations a, semble-t-il, profité davantage aux jeunes issus des quartiers plus défavorisés puisque c'est dans cette catégorie que l'on observe une baisse significative du retard scolaire. Les jeunes issus des quartiers plus favorisés ont maintenu, sur l'ensemble de la période, de bons résultats en termes de scolarités comme le laisse suggérer les faibles taux de retard scolaire.

Le dernier indicateur renseigne sur la part des individus ayant obtenu un diplôme tel que le BEP, le Bac ou un diplôme supérieur (pour les plus avancés) à l'âge de 18 ans et ce, sur l'ensemble de la population considérée. Cet indicateur révèle que l'accès à un diplôme s'est accru au fil du temps. On note une amélioration significative sur la période 1990-2002 de notre mesure de la réussite scolaire (on passe de 68 % à près de 80 % entre ces deux dates). Ce nouveau résultat atteste *in fine* d'une relative démocratisation de l'accès à l'enseignement pour l'ensemble des jeunes.

Par ailleurs, si démocratisation il y a eu, il apparaît que celle-ci a profité à l'ensemble de la population sans trop de disparités. En effet, on observe la même amélioration pour les jeunes issus de quartiers défavorisés ou pour ceux issus de quartiers plus favorisés. Cette amélioration conjointe ne masque pourtant pas une réussite inégale entre les deux populations puisque le taux de réussite est proche de 70 % pour les premiers tandis qu'il atteint presque 80 % pour les seconds.

Ainsi, il semble bien que la réussite scolaire mesurée diffère fortement dès lors que l'on se situe dans une « aire » évaluée comme défavorisée ou dans une « aire » plutôt évaluée comme favorisée. Ces écarts s'expliquent-ils au-delà du contexte familial, par le contexte social et local ? Dans ce travail, nous postulons que le contexte local est un élément clef pour expliquer ces disparités. C'est ce que nous cherchons à montrer dans les sections suivantes.

2.3. La stratégie pour évaluer les effets de voisinage

Notre objectif est de mesurer l'effet du contexte local sur la réussite scolaire des jeunes. Néanmoins, la littérature appliquée à ce sujet montre qu'il existe de potentiels problèmes de biais de sélection. C'est typiquement le cas dès lors que l'on considère que le choix de résider dans tel ou tel quartier n'est pas aléatoire. Afin de prendre en compte ce type de biais, nous avons recours aux méthodes d'évaluation par appariements sélectifs. Initialement développées pour l'évaluation de l'efficacité des traitements dans le domaine médical, ces méthodes ont été importées et appliquées dans le domaine des sciences sociales par Rubin (Rubin, 1974 ; Rubin et Rosenbaum, 1983) dans le cas d'*expériences naturelles* (donc non nécessairement contrôlées).

Rappelons ici que l'idée est de comparer les effets du voisinage sur les performances de deux groupes d'adolescents, toutes choses égales par ailleurs. Ces individus résident soit dans un quartier défini comme défavorisé (individus *traités*), soit, à l'inverse, dans un quartier défini comme plus favorisé (individus *non traités*). Nous expliquons ici de manière intuitive, plus que formelle, la méthode d'appariement sélectif (voir Annexe 1 pour plus de détails). Dans ce travail, nous cherchons à mettre en évidence un effet causal du lieu de résidence sur la réussite scolaire. L'effet causal, ici, est défini comme la différence en termes de résultats entre des individus recevant un traitement donné et le « contrefactuel » : c'est-à-dire, les mêmes individus ne recevant pas ledit traitement. Dans le cas présent, le traitement correspond au fait de vivre dans un voisinage mesuré comme « défavorisé ». Comme un même individu ne peut pas simultanément habiter *et* ne pas habiter dans ce type de voisinage, nous utilisons la méthode d'appariement. L'idée est alors d'apparier chaque individu traité avec un ou plusieurs individus appartenant au groupe de contrôle¹⁰, de telle sorte que les individus traités soient, en moyenne, identiques à ceux du groupe de contrôle. Cet

¹⁰ Le groupe de contrôle constitue par définition le groupe des individus qui n'ont pas bénéficié du traitement.

appariement, qui se fait sur la base des caractéristiques observables, permet d'identifier les individus qui servent de contrôle définissant le « contrefactuel ».

Le groupe de contrôle (ou groupe témoin) et le groupe traité sont bien appariés dès lors que les caractéristiques individuelles affectant le résultat sont « équilibrées » entre les deux. Idéalement, l'appariement se fait sur l'ensemble des caractéristiques individuelles. Dans les faits, ceci est rarement possible en raison de la taille restreinte de certaines bases de données et de l'appariement sur un trop grand nombre de variables. Une stratégie alternative, qui est également la nôtre dans ce travail, est de créer une mesure synthétique de la probabilité de recevoir le traitement¹¹. Concrètement nous calculons la probabilité de vivre dans un quartier où le capital socio-culturel mesuré est faible, grâce à une estimation *logit*, en introduisant un grand nombre de variables susceptibles d'influer sur ce fait. Dès lors, les individus seront comparables s'ils sont proches en termes de probabilité d'accès au traitement.

Finalement, en comparant des jumeaux d'individus dont les caractéristiques observables se rapprochent le plus possible, mais qui se différencient selon le fait qu'ils soient *traités* ou *non traités* (autrement dit s'ils habitent dans un environnement où le capital socio-culturel est faible ou élevé), il devient possible de déterminer l'effet du traitement sur un résultat donné (ici la réussite scolaire). Puisque nous contrôlons de toutes les caractéristiques observables, les différences mesurées entre les deux groupes ne peuvent vraisemblablement s'expliquer que par un effet du traitement.

Néanmoins, évaluer l'impact du lieu de résidence sur la réussite scolaire, dès lors que l'établissement scolaire de l'adolescent ne coïncide pas avec son lieu de résidence, peut poser problème. En effet, si deux enfants d'un même voisinage vont dans des établissements différents, en termes de réussite scolaire, les résultats de nos estimations peuvent vraisemblablement être biaisés. Toutefois, l'existence de la carte scolaire et de la sectorisation permet de s'assurer qu'une grande majorité des adolescents s'inscrivent effectivement dans l'établissement proche de leur domicile¹².

Les modèles d'appariement sur le score de propension que nous mobilisons ici sont donc des méthodes en deux étapes. Après avoir estimé le score de propension des traités et des contrefactuels, à l'aide de notre modèle *logit*, nous procédons à une estimation par appariement sélectif. Il existe plusieurs méthodes dans la littérature permettant l'appariement entre les traités et les non-traités. Nous retenons celle du plus proche voisin (*Nearest Neighbour Matching*).

La méthode du plus proche voisin consiste à choisir pour chaque individu traité le contrefactuel qui est le plus proche en termes de score de propension. L'appariement avec le plus proche voisin peut être effectué avec « remise » (un individu non traité sera potentiellement utilisable plusieurs fois en tant que « contrefactuel ») ou « sans remise » (ici l'individu ne peut être sélectionné qu'une seule fois). Dans le prolongement de cette méthode, il est aussi possible d'utiliser plusieurs plus proches voisins, au lieu d'un seul. Que l'on choisisse le tirage avec « remise » ou l'utilisation de plus qu'un seul plus proche voisin, la difficulté sera d'arbitrer entre variance et biais (Caliendo et Copening, 2005). Dans le premier cas, la « remise » permet d'accroître la qualité des appariements et *in fine* réduit le biais, mais au détriment d'une augmentation de la variance. Dans le deuxième cas, utiliser

¹¹ Cette méthode est plus connue sous le nom d'appariement par le score de propension.

¹² Par exemple, selon une enquête menée en 1998 sous la responsabilité de la Direction de la programmation et du développement du ministère de l'Éducation nationale, seulement 9 % des élèves entrant en sixième intègrent un collège public hors du secteur du collège dont dépend leur habitation.

plusieurs voisins réduit la qualité de l'appariement et donc accroît le biais potentiel, mais le fait de mobiliser plus d'informations pour construire le « contrefactuel » permet de réduire la variance.

Finalement, nous retenons la méthode d'appariement par le plus proche voisin avec « remise » et celles des plus proches voisins pour vérifier si les résultats sont modifiés et pour attester de la robustesse de ces derniers. La première de ces deux méthodes se justifie par la distribution de nos deux populations et du support commun (voir graphique 4). En effet, le support commun entre les individus traités et non traités est relativement faible, il devient alors nécessaire d'accroître les chances de réaliser un bon appariement en autorisant les « remises ». Néanmoins comme cette méthode tend à accroître la variance, nous retenons également la seconde méthode qui, elle, minimise la variance mais au détriment d'une hausse du biais. La section suivante présente les résultats pour ces deux alternatives.

3. RÉSULTATS

Après avoir présenté une estimation par un modèle *logit* de la probabilité de vivre dans un voisinage défavorisé, nous montrons les conséquences de vivre dans un tel quartier (comparativement au fait de vivre dans un quartier favorisé) sur la réussite scolaire des adolescents. Nous présentons donc, dans un second temps, les résultats du modèle d'appariement par la méthode du score de propension.

3.1. Vivre dans un quartier défavorisé : quels sont les déterminants ?

Nous mettons en œuvre une modélisation *logit* pour expliquer la probabilité, pour un adolescent donné, d'habiter dans un voisinage « défavorisé ». Le tableau 3 présente les résultats. En premier lieu, il apparaît que le diplôme ou la catégorie socio-professionnelle du chef de ménage influe sur la probabilité de résider dans un quartier défavorisé. Concrètement, plus le chef de ménage est diplômé, plus l'adolescent a de chances de grandir dans un environnement favorable. Inversement, les quartiers défavorisés abritent davantage de familles dont le chef de ménage est peu ou pas diplômé. De même, les familles dont le chef de ménage appartient à une catégorie socio-professionnelle élevée (cadres, professions libérales, chefs d'entreprises) résident plus souvent dans une aire où le niveau socio-culturel mesuré est important. À l'inverse, être issu d'une famille dont le chef est d'une catégorie socio-professionnelle plus modeste augmente la probabilité de vivre dans un quartier défavorisé.

Concernant les caractéristiques du ménage, d'autres variables jouent également un rôle important. Par exemple, vivre dans une famille dont le chef de famille est une femme augmente les risques de résider dans ces aires défavorisées. Dans la grande majorité des cas (pour 97 % d'entre elles), il s'agit de familles monoparentales. En outre, le nombre d'enfants augmente lui aussi les risques, puisque la probabilité de vivre dans ce type d'aires augmente avec le nombre d'enfants dans le ménage. Nos données montrent également qu'être de nationalité étrangère augmente la probabilité de grandir dans un quartier défavorisé.

Tableau 3 :
Modélisation *logit* de la probabilité d'habiter dans un quartier défavorisé

Variables	Coefficients	écart-type
Caract. individuelles		
Homme	-0,03	0,050
Nationalité française	-0,48***	0,125
Mois de naissance	-0,01	0,007
Caract. du ménage		
vie en couple	0,31**	0,110
Un enfant	-0,16***	0,068
Deux enfants	<i>Réf.</i>	
Trois enfants	0,20***	0,064
Quatre enfants ou plus	0,64***	0,082
Caract. du chef du ménage		
Homme	-0,47***	0,129
En activité	0,01	0,106
- Diplôme		
Diplôme supérieur	-0,72***	0,110
Bac + 2	-0,01	0,119
Bac ou Brevet Pro.	<i>Réf.</i>	
CAP, BEP	0,51***	0,089
BEPC	0,23*	0,117
Aucun diplôme	0,60***	0,092
- Catégorie socio-professionnelle		
Chefs d'entreprises, artisans	-0,93***	0,155
Cadres, prof. Int.	-0,31***	0,087
Employés	<i>Réf.</i>	
Ouvriers	0,20**	0,089
Autres	-0,16	0,145
Caract. du lieu de résidence		
Commune rurale	2,75***	0,121
Unité urbaine < 20 000 hab.	3,02***	0,125
Unité urb. 20 000 à 200 000 hab.	1,86***	0,117
Unité urbaine > 200 000 hab.	-0,11	0,109
Agglomération parisienne	<i>Réf.</i>	
Indicatrices de régions	Oui	
Indicatrices de générations	Oui	
Observations	12 236	
Log Likelihood	-5094,715	
Pseudo-R²	0,3296	

Source : Insee, Enquêtes « Emploi », 1990-2002.

Champ : Ensemble des 12 236 adolescents vivant dans un quartier déterminé comme « défavorisé » ou « favorisé ».

Niveaux de significativité : *** au seuil de 1% ; ** au seuil de 5% ; * au seuil de 10%.

Nous introduisons également un certain nombre de caractéristiques des adolescents qui, bien que peu évidentes à mettre en relation avec la probabilité de vivre dans tel ou tel

quartier, nous semblent néanmoins pertinentes pour avoir des individus comparables¹³. Ainsi, si le sexe ou même le mois de naissance n'expliquent pas le lieu de résidence, ils peuvent potentiellement être des déterminants de la réussite scolaire¹⁴.

Enfin, si l'on s'intéresse aux disparités territoriales, on remarque que les régions sinistrées du nord de la France sont les plus propices à la concentration de quartiers défavorisés. Ainsi, par exemple le fait d'habiter dans la région Nord-Pas-de-Calais augmente la probabilité de résider dans un quartier défavorisé. La région Ile-de-France, quant à elle se situe davantage dans une position intermédiaire. Elle s'avère ni être la région la plus défavorable ni la plus favorable. Parmi les plus favorables justement on retrouve la région Languedoc-Roussillon ou le Limousin. Il ressort également que le fait d'habiter dans une commune rurale ou dans une commune n'appartenant pas à une grande agglomération (supérieure à 200 000 habitants) accroît les risques de résider dans une aire « défavorisée ». L'effet pour la région parisienne est plus ambigu puisqu'il inclut à la fois des communes aisées notamment Paris et la banlieue sud-ouest et d'autres plutôt désavantagées en particulier en Seine-Saint-Denis ou dans une partie du Val-de-Marne.

3.2. Les résultats de l'appariement sur le score de propension

Habiter dans un quartier défavorisé joue-t-il sur la réussite scolaire, toutes choses égales par ailleurs ? C'est à cette question que ce travail tente de répondre. Les résultats sont exposés dans le tableau suivant (tableau 4).

Tableau 4 :
Effet estimé de résider dans un voisinage défavorisé sur la réussite scolaire

Spécification 1 : le plus proche voisin (avec remise)				
Mesures de la réussite scolaire	Naïf	écart-type	Effet moyen global	écart-type
Être encore en cours	-0,028***	0,004	-0,014*	0,008
Avoir un diplôme	-0,119***	0,009	-0,050***	0,017
Être en retard	0,095***	0,007	0,034**	0,014
Spécification 2 : les n plus proches voisins (n = 3)				
Mesures de la réussite scolaire	Naïf	écart-type	Effet moyen global	écart-type
Être encore en cours	-0,028***	0,004	-0,012	0,007
Avoir un diplôme	-0,119***	0,009	-0,049***	0,017
Être en retard	0,095***	0,007	0,032**	0,013

Source : Insee, Enquêtes « Emploi », 1990-2002.

Champ : Ensemble des 12 236 adolescents vivant dans un quartier déterminé comme « défavorisé » ou « favorisé ».

Lecture : L'effet naïf correspond à l'écart de performance moyenne entre les individus issus de quartiers favorisés et défavorisés. Ainsi, les adolescents vivant dans un quartier défavorisé ont 2,8 % de chances de plus que les autres de ne plus être en cours. L'effet moyen global traduit la même idée mais tient compte d'un éventuel biais de sélection.

Niveaux de significativité : *** au seuil de 1% ; ** au seuil de 5% ; * au seuil de 10%.

¹³ Condition indispensable et nécessaire pour que notre évaluation par appariement sur le score de propension soit pertinente.

¹⁴ Voir notamment l'article de Goux et Maurin (2005) dans lequel ils explicitent en quoi le mois de naissance pourrait expliquer la réussite scolaire.

Rappelons que nous comparons ici des individus issus de quartiers « favorisés » et « défavorisés » exclusivement. Nous cherchons des individus semblables entre les deux populations et qui, de ce fait, ne diffèrent que par leur lieu de résidence. On peut voir sur le graphique 4 la distribution des scores de propension parmi les individus habitants dans l'un ou l'autre de ces quartiers¹⁵. Dans l'ensemble, les résultats obtenus par notre méthode d'appariement sélectif semblent confirmer les faits déjà observés au niveau des statistiques descriptives.

Le tableau 4 présente les résultats pour les deux spécifications retenues : le plus proche voisin avec « remise » et les n plus proches voisins. Il présente aussi les résultats pour deux types d'estimateurs : (1) un estimateur naïf qui se définit comme l'écart de performance moyenne entre individus résidant dans un quartier défavorisé et ceux vivant dans un quartier favorisé. Ce premier estimateur consiste à comparer la probabilité de réussite à l'école des individus issus des deux types de quartiers. Toutefois cet estimateur est biaisé, car il ne tient pas compte du fait que les adolescents traités et non traités n'aient pas, globalement, les mêmes caractéristiques. (2) Ainsi, dans un second temps, nous avons recours à un estimateur de l'effet moyen global¹⁶ qui prend en compte l'existence d'un tel biais.

Toutes choses égales par ailleurs, la probabilité d'abandonner ses études ne semble pas fondamentalement plus élevée lorsque l'on habite dans un quartier où le capital socio-culturel mesuré est faible. La différence entre les deux populations est faiblement significative dans notre première spécification et ne l'est plus dans la seconde. En revanche, on note dans l'ensemble des effets très significatifs du lieu de résidence sur la probabilité d'obtenir un diplôme tel que le BEP, le Bac ou un diplôme niveau Bac+2. Les adolescents qui résident dans ces quartiers perçus comme « défavorisés » échouent donc plus souvent l'année suivante¹⁷. Une mesure alternative de la réussite, représentée par notre indicateur « être en retard » (c'est-à-dire ne pas être dans un niveau scolaire qui correspond à l'âge de 18 ans), montre les mêmes effets. Ainsi, habiter dans un voisinage défavorable en termes de caractéristiques socio-culturelles influe positivement sur la probabilité d'être en retard et par voie de conséquence sur l'échec scolaire.

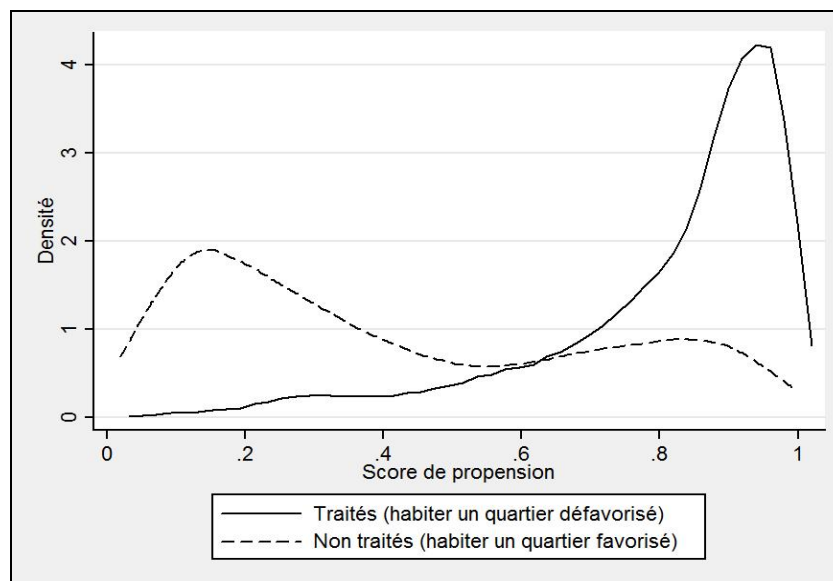
Dans l'ensemble, l'effet naïf est toujours au moins deux fois plus élevé que l'effet obtenu lorsque l'estimateur contrôle du biais de sélection potentiel. Les différences entre ces deux estimateurs montrent la nécessité de contrôler des caractéristiques individuelles lorsque l'on veut comparer la réussite scolaire d'individus habitant dans des quartiers différents. En effet, les résultats plus élevés obtenus avec le premier estimateur révèlent que sans contrôler du biais de sélection, on attribue à tort au lieu de résidence une partie de l'effet des caractéristiques individuelles sur la réussite scolaire. Toutefois, même en ayant contrôlé de ces caractéristiques individuelles, il ressort qu'une partie non négligeable de l'échec scolaire s'explique par des effets de pairs ou de voisinage.

¹⁵ La distribution de ce score sur les groupes de traitement comparés renseigne sur la comparabilité entre ces deux groupes. Plus le support commun entre les deux distributions sera important, moins il sera difficile de trouver des individus comparables en termes de score de propension (ou de probabilité d'être traité).

¹⁶ On utilise dans ce cas l'estimateur par appariement dont la méthodologie est décrite dans la section précédente et en annexe 1.

¹⁷ Rappelons ici que l'on étudie l'impact d'habiter dans un quartier défavorisé pour des jeunes de 17 ans sur leur réussite scolaire mesurée un an plus tard.

Graphique 4 :
Distribution des scores de propension
pour les adolescents traités et non traités



Source : Insee, Enquêtes « Emploi », 1990-2002.

Champ : Ensemble des 12 236 adolescents vivant dans un quartier déterminé comme « défavorisé » ou « favorisé ».

Lecture : La zone du graphique qui est en dessous de la distribution du score de propension, pour les individus traités et en dessous de celle des individus non traités, représente le support commun.

3.3 Sensibilité des résultats à l'hétérogénéité inobservée

Évaluer l'effet d'un traitement par la méthode d'appariement sélectif suppose que l'hypothèse d'indépendance conditionnelle soit vérifiée. Or, si les deux groupes (traités et non-traités) diffèrent sur des variables inobservées qui affectent simultanément la probabilité d'être traité et la variable de résultat, alors il peut exister un « biais caché ». L'estimation par appariement, dans son application actuelle, ne permet pas la prise en compte d'un tel biais.

Concrètement, dans notre cas, lorsque nous essayons de mettre au jour les effets du lieu de résidence sur la réussite scolaire, il est probable que certaines variables influent sur les deux sans que nous puissions contrôler de cet effet. En effet, outre les caractéristiques individuelles et familiales (qui font parties des caractéristiques observables et qui sont contrôlées dans l'analyse) qui peuvent expliquer la réussite scolaire, il peut également exister un effet « établissement ». Ainsi, la différence de réussite scolaire entre les individus « traités » et « non-traités » peut effectivement s'expliquer par des effets de voisinage mais aussi par le fait que les établissements auxquels ils sont affiliés diffèrent en termes de performances. Les ménages pourraient alors choisir leur lieu de résidence en fonction des résultats et de la renommée d'un établissement donné. On a bien, dans ce cas, un variable inobservable¹⁸ qui peut affecter la probabilité d'être « traité » ou non et qui influe sur la réussite scolaire.

¹⁸ L'enquête emploi ne nous permet pas d'avoir des informations sur l'établissement dans lequel les adolescents suivent leur cursus scolaire.

Afin de déterminer à quel point une variable non-mesurée peut influencer sur le processus de sélection, rendant les résultats de nos estimations par appariements sélectifs non significatifs, nous mobilisons la méthode d'analyse de sensibilité de Rosenbaum (2002). Nous décrivons brièvement la méthodologie adoptée dans le paragraphe suivant.

La probabilité d'être « traité » pour un individu i avec les caractéristiques observables x_i peut être écrite de la manière suivante : $P(x_i) = P(D_i = 1 | x_i) = F(\beta x_i + \gamma u_i)$ où u_i représente la variable inobservée et γ représente l'effet de u_i sur la probabilité d'être « traité ». Si il n'existe pas de « biais caché » alors γ sera égal à 0 et la probabilité d'être « traité » dépendra exclusivement de x_i . Cependant, si ce biais existe, alors des individus avec des caractéristiques observables similaires afficheront différentes probabilités de recevoir le traitement. Suivant Aakvik (2001), qui propose de simplifier l'analyse en traitant la variable inobservée comme une variable dichotomique, prenant la valeur 0 (en l'absence de biais) ou la valeur 1 (en présence de biais), la variable e^γ peut être vue comme une mesure de l'écart par rapport à une situation exempte de tout biais. Comme le montre Rosenbaum (2002), le rapport des chances (ou odds-ratio) que l'un des deux individus appariés reçoive le traitement a les bornes suivantes :

$$\frac{1}{e^\gamma} \leq \frac{P(x_i)(1 - P(x_j))}{P(x_j)(1 - P(x_i))} \leq e^\gamma$$

Les individus appariés ont la même probabilité d'être traités si et seulement si $e^\gamma = 1$. Dès lors, l'idée est d'augmenter l'influence de e^γ et de vérifier si les conclusions de l'analyse statistique peuvent être invalidées. Pour cela, nous utilisons le test statistique de Mantel et Haenszel (1959). Ce test permet de vérifier l'hypothèse nulle d'indépendance entre deux variables dichotomiques (dans notre cas « être traité » et « réussir à l'école ») sur une population divisée en plusieurs classes (ici, on peut imaginer deux classes : être dans un bon établissement ou non). Ce test suppose que toute association entre les variables dichotomiques n'est pas affectée par la troisième variable.

Le tableau 5 présente les résultats de notre analyse de sensibilité. L'effet négatif du fait d'appartenir à un voisinage défavorisé sur le fait d'être encore en cours, trouvé dans la section précédente, apparaît être le moins robuste à la présence éventuelle d'un biais de sélection. En effet, dans ce cas là, le seuil critique à partir duquel les conclusions sur un effet négatif deviennent questionnables est atteint pour une valeur de e^γ entre 1,10 et 1,15. En d'autres termes, l'effet négatif mis en avant peut être remis en question dès lors qu'une variable inobservée fait varier le rapport des chances d'être traité entre les deux groupes par un facteur d'environ 1,15. Concernant les deux autres indicateurs testés : « être diplômé » et « être en retard scolaire », l'ampleur du « biais caché » devrait être respectivement de 1,25 et 1,4 afin de remettre en question l'effet de résider dans un voisinage défavorisé sur la réussite scolaire.

Cette variable inobservée qui peut effectivement biaiser les résultats de notre analyse et remettre en cause la robustesse des résultats pourrait ainsi s'identifier à cet effet « établissement ». En effet, si l'on part du principe que les ménages choisissent leur lieu de résidence pour bénéficier d'une sectorisation favorable à l'éducation de leurs enfants, alors l'hypothèse d'indépendance entre le lieu de résidence et la réussite scolaire ne tient plus. Toutefois, les résultats de l'analyse de sensibilité révèlent que la robustesse n'est questionnable qu'au-delà d'un certain seuil. Ils montrent également que ce processus de sélection tend à surestimer l'effet du traitement (*i.e.* habiter dans un quartier défavorisé) sur la réussite scolaire. Ainsi, une partie de l'échec scolaire est vraisemblablement attribuée à

tort à l'effet du voisinage, négligeant ainsi cet effet « établissement ». Inversement, l'effet du traitement sur l'échec scolaire (ou être en retard scolaire) tend, quant à lui, à être sous-estimé. L'établissement dans lequel est scolarisé influe donc sur la réussite individuelle sans que cet effet soit véritablement capturé, cependant les résultats ne doivent pas forcément être remis en cause puisque la non-significativité n'intervient qu'au-delà d'un certain seuil.

Tableau 5 :
Analyse de sensibilité pour l'hétérogénéité inobservée¹⁹

Spécifications	Être en cours	Être diplômé	Être en retard scolaire
e^γ	p-MH+	p-MH+	p-MH-
1	0,029	> 0,001	> 0,001
1,05	0,061	> 0,001	> 0,001
1,1	0,102	> 0,001	> 0,001
1,15	0,184	0,007	> 0,001
1,2	0,275	0,035	> 0,001
1,25	0,379	0,117	0,004
1,3	0,489	0,278	0,015
1,35	0,461	0,496	0,044
1,4	0,362	0,316	0,103
1,45	0,274	0,155	0,201
1,5	0,201	0,062	0,335

Source : Insee, Enquêtes « Emploi », 1990-2002.

Lecture : p-MH représente le niveau de significativité pour la statistique de Mantel-Haenszel. Pour p-MH+, l'hypothèse testée est la sur-estimation de l'effet du traitement. Pour p-MH-, l'hypothèse testée est la sous-estimation de l'effet du traitement.

Enfin, il importe de préciser que les seuils de 1.15, 1.20 et 1.4 représentent les pires scénarios. La valeur de 1,15 ne signifie pas forcément que l'hétérogénéité inobservée existe et qu'il n'y a pas d'effet du lieu de résidence sur le fait d'être en cours ou non. Cela signifie que l'intervalle de confiance pour cet effet inclurait 0 si une variable inobservée faisait varier le rapport des chances d'être « traité » entre les deux groupes d'un facteur 1,15. Puisque les résultats apparaissent sensibles à un écart par rapport à une situation où l'hypothèse d'indépendance conditionnelle est vérifiée, leur interprétation doit donc être faite avec précaution.

CONCLUSION

L'objectif de cet article est de montrer qu'outre le niveau socio-culturel des parents, d'autres déterminants sont à considérer pour expliquer les écarts de réussite scolaire observés entre les adolescents en France. Dans notre cas, on se concentre en particulier sur les effets de voisinage et les conséquences que ceux-ci peuvent induire sur la réussite scolaire des jeunes. La question de la nécessité d'une forme de mixité sociale dans les quartiers dits

¹⁹ Les seuils de significativité pour l'hypothèse de sous-estimation (p-MH-) pour « être en cours » et « être diplômé » ainsi que pour l'hypothèse de sur-estimation (p-MH+) pour « être en retard » ne sont pas présentés car toujours significatifs.

défavorisés est ainsi soulevée. L'idée générale de ce travail est donc de voir si, toutes choses égales par ailleurs, habiter dans un quartier perçu comme « défavorisé » a un effet négatif sur la réussite scolaire. Ce faisant, un problème de biais de sélection semble se poser puisqu'il est vraisemblable que le choix du lieu de résidence des ménages ne s'effectue pas de manière aléatoire. Pour corriger de ce biais, nous avons recours aux méthodes d'évaluation par appariements sélectifs.

Finalement, nos résultats montrent qu'il existe un effet du contexte social sur la scolarité, au-delà de l'influence des caractéristiques individuelles et familiales du jeune qui reste primordiale pour expliquer leur réussite scolaire. En outre, si notre travail permet justement de contrôler des caractéristiques individuelles et familiales qui jouent un rôle important, il ne permet pas de contrôler d'un éventuel effet « établissement ». Cette variable inobservable pourrait biaiser l'analyse dans le sens où elle peut être à l'origine d'un processus de sélection. En effet, les ménages pourraient privilégier les quartiers proches d'établissements à bonne réputation, ce qui influencerait sur la réussite scolaire mesurée. Pour évaluer l'ampleur et la pertinence d'un tel biais, nous avons eu recours à une analyse de sensibilité des résultats à l'hétérogénéité inobservée. Sans remettre en cause les conclusions de ce travail, l'analyse montre qu'il faut être prudent dans l'interprétation des résultats.

Par ailleurs, bien que notre méthode ne permette pas de dire lequel des mécanismes évoqués prévalent (effets de contagion, problème de frein à l'accumulation de capital humain...), il demeure toutefois possible d'en tirer des conclusions en termes de politiques publiques. On pense, entre autres, à bon nombre de programmes qui développeraient ou encourageraient la mixité sociale au sein des quartiers, et notamment au sein des institutions scolaires. Une autre alternative résiderait dans la mise en place de politiques ciblées notamment sur les familles les plus exposées à l'échec scolaire, et donc vers les familles qui connaissent de grandes difficultés sociales. Avec l'idée sous-jacente selon laquelle améliorer la performance de quelques-uns peut vraisemblablement améliorer la performance de tous. Et ce, par les mêmes mécanismes qui contribuent à freiner la réussite scolaire des adolescents d'un quartier : C'est-à-dire les mouvements de reflets ou ce que l'on pourrait appeler l'influence respectueuse. D'où, la nécessité de déployer des moyens (budgétaires et en personnel) importants dans les institutions scolaires qui se trouvent dans des quartiers défavorisés pour faire face à une concentration d'élèves en difficultés.

BIBLIOGRAPHIE

- AAKVIK A. (2001), "Bounding a matching estimator: The case of a Norwegian training program", *Oxford Bulletin of Economics and Statistics*, vol. 63, pp. 115-143.
- BECKER S., CALIENDO M. (2007), "Sensitivity analysis for average treatment effects", *The Stata Journal*, vol. 7 (1), pp. 71-83.
- BENABOU R. (1993), "Working of a city : Location, education and production", *Quarterly Journal of Economic*, vol. 108, pp. 619-652.
- BRODATY T., CREPON B. et FOUGERE D. (2007), "Les méthodes micro-économétriques d'évaluation et leurs applications aux politiques actives de l'emploi", *Économie et Prévision*, n°177, pp. 93-118.
- BRODATY T., GURGAND M. (2008), "Teacher and peer effects in higher education: Evidence from a French university", *Working Paper*.
- CALIENDO M., KOPEINIG S. (2005), "Some practical guidance for the implementation of propensity score matching", *IZA Discussion Paper*, n°1588.

- CRANE J. (1991), " The epidemic theory of ghettos and neighbourhood effects on dropping out and teenage childbearing ", *American Journal of Sociology*, vol. 96, pp. 1226-1259.
- CUTLER D., GLAESER E. (1997), " Are ghettos good or bad ? ", *Quarterly Journal of Economics*, vol. 112, pp. 827-872.
- DIETZ R. (2002), " The estimation of neighborhood effects in the social sciences: An interdisciplinary approach ", *Social Science Research*, vol.31, pp. 539-575.
- DIPRETE T., GANGL, M. (2004), " Assessing bias in the estimation of causal effects : Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments ", *Sociological Methodology*, vol. 34, pp. 271-310.
- GINTHER D., HAVEMAN R. et WOLFE B. (2000), " Neighborhood attributes as determinants of children's outcomes: How robust are the relationships ? ", *The Journal of Human Resources*, vol. 35 (4), pp. 603-642.
- GOUX D., MAURIN E. (2005), " Composition sociale du voisinage et échec scolaire ", *Revue Economique*, vol. 56 (2), pp. 349-362.
- HARDING D. (2003), " Counterfactual models of neighborhood Effects: The effect of neighborhood poverty on dropping out and teenage pregnancy", *American Journal of Sociology*, vol. 109 (3), pp. 676-719.
- JENCKS C., MAYER E. (1990), " The social consequences of growing up in a poor neighbourhood ", *Inner-City Poverty in the United States*, LYNN L. E. et MCCOY M. G., Washington, (D.C.), National Academy Press..
- MANSKI C. (1993), " Identification of endogenous social effects : the reflection problem ", *Review of Economic Studies*, vol. 60.
- MANSKI C. (2000), " Economic analysis of social interactions", *Journal of Economic Perspectives*, vol. 14, pp. 115-136.
- MANTEL N., HAENSZEL W. (1959), " Statistical aspects of the analysis of data from retrospective studies ", *Journal of the National Cancer Institute*, vol 22., pp. 719-748.
- MARPSAT M. (1999), " La modélisation des "effets de quartier" aux États-Unis ", *Population*, vol. 54 (2), pp. 303-330.
- PIKETTY T. (2004), " L'impact de la taille des classes et de la ségrégation sociale sur la réussite scolaire dans les écoles françaises : une estimation à partir du panel primaire 1997 ", *Document de travail*.
- ROSENBAUM P. (2002), *Observational Studies*, 2nde édition. New York : Springer.
- ROSENBAUM P., RUBIN D. (1983), " The Central Role of the Propensity Score in Observational Studies for Causal Effects ", *Biometrika*, vol.70 (1), pp. 41-55.
- RUBIN D. (1974), " Estimating causal effects of treatments in randomized and nonrandomized Studies ", *Journal of Educational Psychology*, vol. 66 (5), pp. 688-701.
- SAMPSON R., MORENOFF J. et GANNON-ROWLEY T. (2002), " Assessing 'neighborhood effect': Social processes and new directions in research ", *Annual Review of Sociology*, vol. 28, pp. 443-478.
- WILSON W. (1987), *The truly disadvantaged: the inner city, the underclass and public policy*, Chicago: University of Chicago Press.

ANNEXE 1 : L'ÉVALUATION DES EFFETS DE QUARTIER PAR LA MÉTHODE D'ÉVALUATION PAR APPARIEMENTS SÉLECTIFS

Notre objectif est de mesurer l'effet du contexte local sur le développement des enfants. Rappelons ici que l'idée est de comparer les effets du voisinage sur les performances de deux groupes d'adolescents, toutes choses égales par ailleurs. L'un vivant dans un voisinage où le niveau socioculturel mesuré est faible et l'autre où le niveau socioculturel est plus important.

Ainsi, en notant D_i le lieu de résidence d'un adolescent donné, on peut considérer que $D_i = 1$ pour les adolescents qui résident dans un quartier défini comme défavorisé, et $D_i = 0$ pour les adolescents qui résident, à l'inverse, dans un quartier défini comme plus favorisé. Dans notre cas, les individus vivent soit dans un environnement où le capital socio-culturel est faible : individus « traités » ($D_i = 1$), soit dans un environnement où le capital socio-culturel est plus élevé : individus « non-traités » ($D_i = 0$).

Si maintenant Y_i représente la réussite scolaire (avoir abandonné ses études ou non, par exemple), alors l'effet contextuel ou l'effet de voisinage sur les comportements des adolescents et leur réussite peut s'écrire :

$$R_i = Y_i(1) - Y_i(0) \quad (1)$$

Comme il est impossible d'observer un même individu face aux deux situations de traitement, on mesure l'effet moyen du traitement sur la population traitée (dans notre cas il s'agit de la population qui habite dans un quartier défavorisé) que l'on note :

$$R_1 = E[Y(1) - Y(0)|D=1] = E[Y(1)|D=1] - E[Y(0)|D=1] \quad (2)$$

La deuxième partie de l'égalité (2) ne peut pas être observée. Pour mesurer l'effet moyen du traitement sur la population traitée, il est nécessaire de supposer que $E[Y(0)|D=1] = E[Y(0)|D=0] = E(Y(0))$, ce qui revient à supposer que quelque soit le milieu socio-culturel dans lequel on réside (ou que l'on côtoie), le comportement à l'école reste le même. On a alors :

$$R_1 = E[Y(1)|D=1] - E[Y(0)|D=0] \quad (3)$$

Par ailleurs, si les populations « traitées » et « non-traitées » ont des caractéristiques différentes, et que celles-ci ont une influence sur Y_i (la réussite scolaire), les résultats de l'estimation de l'effet moyen du traitement peuvent être biaisés. Il convient alors de corriger ce biais en contrôlant des différences de caractéristiques observables X . Le traitement est alors considéré comme distribué aléatoirement « conditionnellement aux caractéristiques observables ».

$$Y(1), Y(0) \perp D | X \quad (4)$$

Quand le nombre de ces caractéristiques est élevé, il est difficile de réaliser un appariement sur l'ensemble de ces caractéristiques. Une solution est de réaliser un appariement sur le score de propension $\pi(x)$, qui se définit comme la probabilité, pour un individu possédant les caractéristiques x , d'être traité ($\pi(x) = \Pr(T=1|X)$). On peut alors écrire :

$$Y(1), Y(0) \perp T \mid \pi(x) \quad (5)$$

L'effet causal est évalué comme :

$$R_i = E_{\pi(X)} (E[Y(1)|D=1, \pi(X)] - E[Y(0)|D=0, \pi(X)] \mid T=1) \quad (6)$$

Il est toutefois essentiel qu'il existe un intervalle commun aux deux distributions du score de propension entre les deux groupes. Soit la condition de « support commun » suivante qui doit être remplie :

$$0 < \pi(X) < 1$$

DERNIERS NUMÉROS PARUS :

téléchargeables à partir du site <http://www.cee-recherche.fr>

- N° 124** *A Dynamic Overview of Socio-Productive Models in France (1992-2004)*
Un panorama dynamique des modèles socio-productifs en France (1992-2004)
THOMAS AMOSSE, THOMAS COUTROT
juin 2010
- N° 123** *Ségrégation urbaine et accès à l'emploi : une introduction*
MANON DOMINGUES DOS SANTOS, YANNICK L'HORTY, ÉLISABETH TOVAR
novembre 2009
- N° 122** *Les effets des allègements de cotisations sociales sur l'emploi et les salaires : une évaluation de la réforme Fillon de 2003*
MATTHIEU BUNEL, FABRICE GILLES, YANNICK L'HORTY
août 2009
- N° 121** *Do Environmental-Related Standards Contribute to Successful Recruitment?*
GILLES GROLLEAU, NAOUFEL MZOUGH, SANJA PEKOVIC
août 2009
- N° 120** *Santé et pénibilité en fin de vie active : Une comparaison européenne*
CATHERINE POLLAK
juin 2009
- N° 119** *Expérimenter pour décider ? Le RSA en débat*
BERNARD GOMEL, EVELYNE SERVERIN
juin 2009
- N° 118** *Réformer les aides sociales locales dans le nouveau contexte du RSA*
DENIS ANNE, YANNICK L'HORTY
mai 2009
- N° 117** *Dépendance interentreprises et inégalités d'emploi : Hypothèses théoriques et tests empiriques*
CORINNE PERRAUDIN, HELOÏSE PETIT, NADINE THEVENOT, BRUNO TINEL, JULIE VALENTIN
mars 2009
- N° 116** *Mesurer la pauvreté et la ségrégation en Île-de-France : une approche capabiliste*
ÉLISABETH TOVAR
mars 2009
- N° 115** *Case Management Services for Jobseekers. International comparisons: Sweden, the Netherlands and the United Kingdom*
NATHALIE GEORGES, NICOLAS GRIVEL, DOMINIQUE MEDA
mars 2009