



HAL
open science

School Resources and Individual Responses : essays in Economics of Education

Manon Garrouste

► **To cite this version:**

Manon Garrouste. School Resources and Individual Responses : essays in Economics of Education. Economics and Finance. Université Panthéon-Sorbonne - Paris I, 2015. English. NNT : 2015PA010010 . tel-01661388

HAL Id: tel-01661388

<https://theses.hal.science/tel-01661388>

Submitted on 11 Dec 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.



École doctorale d'économie Panthéon-Sorbonne n° 465
Université Paris 1 Panthéon-Sorbonne

T H È S E

pour obtenir le grade de docteur

en sciences économiques

présentée et soutenue publiquement par

Manon GARROUSTE

le 30 novembre 2015

**Ressources scolaires et réactions individuelles
Trois essais en économie de l'éducation**

~ ~ ~

Directeur de thèse : **M. Francis KRAMARZ**
Co-directeur de thèse : **M. Nicolas JACQUEMET**

Jury

M. Laurent GOBILLON, Directeur de recherche, CNRS

Mme Miren LAFOURCADE, Professeur d'économie, Université Paris-Sud Rapporteur

M. Eric MAURIN, Directeur d'études, EHESS

M. Alain TRANNOY, Directeur d'études, EHESS Rapporteur

Centre d'économie de la Sorbonne

Remerciements

Avant de commencer, je voudrais dire merci à tous ceux qui m'ont apporté leur aide et leur soutien au cours de cette thèse. Je tiens tout d'abord à remercier Francis Kramarz d'avoir proposé de diriger ma thèse. Je lui suis reconnaissante d'avoir compris mes motivations et mes centres d'intérêt. J'ai apprécié son aide et sa confiance dans mon travail, mais aussi l'indépendance qu'il m'a laissée dans le choix de mes sujets d'étude. Je le remercie également de m'avoir constamment soutenue, et de continuer à le faire, à présent que cette thèse se termine. Je veux dire un immense merci à Nicolas Jacquemet, qui a accepté de co-encadrer mon travail. Je lui suis extrêmement reconnaissante pour son investissement dans cette thèse. Il a toujours lu les travaux que je lui ai envoyés et m'a impressionnée par ses capacités d'écoute et de concentration, et par ses remarques toujours très justes. Je le remercie enfin pour ses encouragements si pertinents : c'est « en se prenant des murs sur le nez » qu'on apprend.

Je remercie ensuite Miren Lafourcade et Alain Trannoy d'avoir accepté d'être rapporteurs. Un grand merci également à Laurent Gobillon et à Eric Maurin qui ont accepté de faire partie du jury. Merci à tous pour leur aide précieuse et leurs remarques. C'est avec plaisir et intérêt que j'ai retravaillé mon manuscrit à la lumière de leurs commentaires. J'espère avoir répondu, au moins en partie, à leurs questions.

Je voudrais remercier chaleureusement tous mes co-auteurs. J'ai pris un immense plaisir à travailler avec eux, et j'espère à la fois continuer à le faire, et rencontrer de nouveaux coéquipiers aussi formidables. Pour commencer, je tiens à remercier Laurent Davezies de m'avoir donné la possibilité de travailler avec lui. J'ai apprécié sa grande rigueur scientifique et sa franchise intellectuelle. Ce sont de belles et trop rares qualités. Je veux dire un grand merci à Jean-Paul Caille pour son aide, dès mes tous premiers travaux à l'Ensaë, et son soutien sans faille depuis. Ses immenses connaissances du système éducatif sont une source infinie d'inspiration. Je le remercie

d'avoir cru en moi et de m'avoir toujours encouragée. Je veux également remercier Ronan Le Saout de m'avoir proposé de travailler avec lui. Merci à Thierry Ly pour son enthousiasme. Enfin, je remercie tout spécialement Meryam Zaiem pour sa spontanéité et son entrain. Travailler avec elle est particulièrement stimulant.

Je voudrais saluer tous mes collègues du Crest. D'abord, je veux dire merci à Philippe Zamora et Roland Rathelot de m'y avoir accueillie. Ensuite, je remercie tout particulièrement Xavier D'Haultfoeuille d'être toujours attentif et disponible. Il contribue à faire du laboratoire de microéconométrie un environnement sain et stimulant. Je remercie énormément Julie pour sa gentillesse et sa détermination, et Marianne pour sa patience et son attention. Je salue bien sûr mes collègues de bureau de la première heure : Maxime, Clément et Benoît. Merci à Mathilde et Pauline pour le chemin fait ensemble. Pour leur aide et leur soutien, merci à Mathias, Léa, Modibo, Diana, Elie, Malka, Audrey, Jérémy, Clémence, Isabelle, Robert, Thierry, Dominique, Sébastien, Sophie, Nadine, Fanda, Arnaud, Edith, Murielle, Martine, Manasa, Bruno, Elise, Bertrand, Yvan, Anna, Thomas, Arne, Edouard, Victor, Anett, Arnaud, Sylvie, Alicia, Mélina, Daphné.

Je remercie mes collègues du centre d'économie de la Sorbonne, notamment Claire, Clémence, Juliette, Léontine, Lorenzo, Théodora et Thomas, d'avoir partagé leurs bureaux et leurs repas avec moi. Je remercie également Elda André et Loïc Sorel pour leur aide et leur efficacité.

Pour leur accueil chaleureux au cours de ma quatrième année de thèse, merci aux membres de l'unité de démographie économique de l'Ined : Anne, Ariane, Bénédicte, Benoît, Carole, Delphine, Dominique, François-Charles, Marion, Olivier, Sébastien. Je remercie en particulier Laurent Gobillon d'avoir accepté d'être mon encadrant à l'Ined. Ses remarques m'ont beaucoup aidée à construire mon projet de recherche. Merci également à Elise, Fabrice, Joanie, Lidia et Xavier. J'ai été très touchée par leur engagement auprès des doctorants. Merci évidemment à tous les doctorants de l'Ined : Abida, Alice, Amélie, Anne, Arlette, Celio, Hélène, John, Maroussia, Meoïn, Milan, Mona, Morgan, Noémie, Paola, Paulina, Pauline, Pierre-Antoine, Sandrine, Valentine, Zoé. Votre solidarité est une force. Valeria, merci pour ta joie de vivre. On pense à toi...

Pour leur accueil, leur aide et leur soutien dans toutes mes démarches auprès de la Direction de l'évaluation, de la prospective et de la performance, je remercie énormément Caroline Simoni-Sueur, Cédric Afsa, Fabrice Murat, Olivier Monso,

Sylvie Le Laidier et Agnès Brizard.

Même s'il serait trop long de les nommer tous, je tiens à remercier tout particulièrement mes étudiants de l'Ensaë, de l'ENPC et de PSE. Ils sont la raison pour laquelle j'ai commencé cette thèse.

Pour leurs encouragements et leur amitié, je remercie Claire, Hélène, Sofia et Tahiry, ainsi qu'Aude, Layla, Geoffrey, Rachid, Sébastien et Vincent. Merci à Miléna, et courage! Merci également à Coralie de m'avoir très justement rappelé que, dans la recherche comme dans le sport, l'important c'est de participer.

Je remercie mes parents de m'avoir donné, chacun à leur manière, l'un pour l'économie, l'autre pour l'éducation, l'envie de faire cette thèse et de continuer. Merci à Clémentine de m'avoir toujours montré la voie et à Jeanne d'avoir suivi la sienne. Je remercie Georges et Cécile pour leur soutien constant durant mes études. J'ai une pensée pour tous ceux qui m'ont accueillie durant la rédaction, notamment Claudine et Jean-Raymond.

Enfin, merci à Bertrand de m'avoir soutenue, aidée et écoutée. Merci pour tout.

Table des matières

Introduction générale	xiii
1 Sorting effects	1
1.1 A brief description of the French education system	5
1.2 Theoretical mechanisms	9
1.3 The data	10
1.4 Effect on individual sorting	16
1.4.1 Empirical strategy	16
1.4.2 Results	23
1.5 Effect on individual academic achievement	32
1.5.1 Empirical strategy	32
1.5.2 Results	35
1.6 Conclusion	47
Appendix	48
1.A Manipulation of the forcing variables	48
1.B Placebo tests	50
1.C Mean comparison of pupils' municipality characteristics	53
1.D Results around discontinuity d^F	56
1.E Estimations on grades	58
1.F Heterogeneous effects on achievement	60
2 High school openings	65
2.1 Institutional Context	67
2.1.1 Track choice at the end of 9th grade	67
2.1.2 Building new high schools	69
2.2 Data	70

2.3	Descriptive Statistics	71
2.4	Estimation strategy	76
2.5	Results	81
2.6	Robustness	89
2.7	Conclusion	90
	Appendix	91
3	Student evaluation of teaching	101
3.1	The data	103
3.2	Descriptive statistics	107
3.3	Empirical strategy and results	111
3.3.1	Empirical strategy	111
3.3.2	Results	113
3.3.3	The dynamics of evaluations over time	117
3.4	Conclusion	121
	Appendix	123
	Conclusion	133

Table des figures

1	La fonction de production de l'éducation	xiv
2	Schéma chapitre 1	xxv
3	Schéma chapitre 2	xxvii
4	Schéma chapitre 3	xxviii
1.1	Map of public junior high schools in the sample	8
1.2	Individual probability that the nearest junior high school is RAR	18
1.3	Illustration of a window around each discontinuity	19
1.4	Individual probability to be enrolled in a RAR junior high school	33
1.5	McCrary Density Test	49
2.1	Illustration of the French school system	68
2.2	Tracks followed by students the year after 9th grade	72
2.3	High school openings in mainland France (2003-2010)	74
2.4	Share of pupils entering a new high school	76
2.5	Evolution of the proportion of pupils allocated to the general track	79
2.6	Evolution of the composition of treated schools before and after an opening	80
3.1	Estimation of the relationship between evaluations and date	120
3.2	Dates of evaluations	130
3.3	Linear relationship between evaluations and exam grade	131
3.4	Multivariate relationship between evaluations and exam grade	132

Liste des tableaux

1.1	Individual characteristics in the sample	13
1.2	Individual characteristics in the sample	14
1.3	Characteristics of pupils' municipality of residence in the sample	15
1.4	Individual characteristics around discontinuity d^L	21
1.5	Individual characteristics around discontinuity d^F	22
1.6	Estimation of the effect of living near a RAR junior high school on school choice	27
1.7	Estimation of heterogeneous effects of living near a RAR on school choice (d^L)	28
1.8	Estimation of the effects of living near a RAR on school choice according to knowledge of the educational system (d^L)	29
1.9	Estimation of the effects of living near a RAR on school choice according to private school supply (d^L)	31
1.10	Estimation of the effect of enrollment in a RAR on passing the <i>Brevet</i> - Discontinuity d10	37
1.11	Estimation of the effect of enrollment in a RAR on passing the <i>Brevet</i> - Discontinuity d67	38
1.12	Estimation of the effect of enrollment in a RAR on the situation 5 years later - Discontinuity d10	40
1.13	Estimation of the effect of enrollment in a RAR on the situation 5 years later - Discontinuity d67	41
1.14	Estimation of heterogeneous effects of enrollment in a RAR on passing the <i>Brevet</i>	43
1.15	Estimation of heterogeneous effects of enrollment in a RAR on Franch grade	44

1.16	Estimation of heterogeneous effects of enrollment in a RAR on enrollment in general high school	45
1.17	Estimation of heterogeneous effects of enrollment in a RAR on repetition	46
1.18	Placebo estimation of the effect of living near a RAR	51
1.19	Placebo estimation of the effect of enrollment in a RAR	52
1.20	Characteristics of pupils' municipality of residence around discontinuity d^L	54
1.21	Characteristics of pupils' municipality of residence around discontinuity d^F	55
1.22	Estimation of heterogeneous effects of living near a RAR on school choice around discontinuity d^F	57
1.23	Estimation of the effect of enrollment in a RAR on <i>Brevet</i> grades - Discontinuity d10	58
1.24	Estimation of the effect of enrollment in a RAR on <i>Brevet</i> grades - Discontinuity d67	59
1.25	Estimation of heterogeneous effects of enrollment in a RAR on passing the <i>Brevet</i>	61
1.26	Estimation of heterogeneous effects of enrollment in a RAR on Math grade	62
1.27	Estimation of heterogeneous effects of enrollment in a RAR on French grade	63
1.28	Estimation of heterogeneous effects of enrollment in a RAR on enrollment in general high school	64
2.1	High schools openings in the sample by year and type	73
2.2	Types of municipalities where high schools are located	75
2.3	Treated middle schools in the sample by year and definition of treatment	75
2.4	DID estimates of the effect of opening a new high school on track choice	85
2.5	DID estimates of the short- and long-term effect of opening a new high school	86
2.6	DID estimates of the effect of opening a new high school depending on local private supply	87
2.7	Separate estimates of the effect of opening a new high school for Ile de France	88
2.8	DID estimates of the effect of opening a new high school by sector	92

2.9	DID estimates of the effect of opening a new high school on the two closest schools	93
2.10	DID estimates of the effect of opening a new high school on the five closest schools	94
2.11	DID estimates of the effect of opening a new high school within median radius	95
2.12	DID estimates of the effect of opening a new high school without "super" control group	96
2.13	DID estimates of the effect of opening a new high school with pre-treatment trend	97
2.14	DID estimates of the effect of opening a new high school on school social composition	98
2.15	DID estimates of the Placebo effect of opening a new high school two years before	99
3.1	OLS estimations of potential determinants of teacher's pedagogy mean evaluation	110
3.2	Estimation of the effect of mean grade on teacher's pedagogy mean evaluation	116
3.3	First stages	117
3.4	Description of the evaluation form	123
3.5	Distribution of evaluations in the sample	124
3.6	Main characteristics of exam grades in the sample	124
3.7	Main characteristics of evaluations in the sample	125
3.8	Correlations between mean evaluations and mean grades	126
3.9	Variance decomposition of evaluations and grades	126
3.10	Estimation of the effect of mean grade on teacher's pedagogy mean evaluation - IV with a change of teacher	127
3.11	First stages with a change of teacher	128
3.12	Estimation of the effect of mean grade on teaching assistant's pedagogy mean evaluation	129

Introduction générale

L'importance des ressources scolaires dans la production de l'éducation

La théorie initiale du capital humain (BECKER, 1964; MINCER, 1974; SCHULTZ, 1961) considère l'acquisition de savoirs et de compétences comme le fruit d'un investissement rationnel de la part des individus pour constituer un capital, dont les rendements futurs permettent d'augmenter leur productivité et donc leurs revenus du travail. Par analogie avec le processus de production d'une entreprise, la transformation des ressources éducatives en résultats scolaires (et *in fine* en salaire) est modélisée sous forme d'une fonction de production de l'éducation. Cette fonction représente la relation entre des « facteurs de production » et la quantité de capital humain produit. Autrement dit, il existe une relation de cause à effet entre les ressources dont l'individu dispose (enseignement scolaire, livres, ressources économiques, etc.) et les connaissances qu'il acquiert. Ici, les facteurs de production sont l'ensemble des ressources éducatives mobilisées pour accumuler du capital humain¹. La quantité de capital humain produite grâce à ces ressources est bien évidemment impossible à mesurer directement. On considère généralement les résultats scolaires (les notes, les qualifications et les diplômes obtenus) et le revenu du travail comme des mesures indirectes du capital humain.

Pour simplifier, on peut considérer qu'il existe deux grands types de ressources (voir schéma figure 1); celles qui sont fournies par l'institution scolaire ou l'école au sens large (les ressources pédagogiques, les enseignants, les moyens fournis par

¹On se restreint ici aux ressources éducatives, bien que G. S. Becker considère également plus largement les dépenses de santé ou l'activité physique par exemple comme autant de façons d'investir dans le capital humain : « The many ways to invest include schooling, on-the-job training, medical care, vitamin consumption, and acquiring information about the economic system. », p.9, BECKER (1962)

l'école, la structure des classes, etc.) et celles qui sont apportées plus directement par l'individu et son environnement familial : le capital économique et culturel des parents, les capacités cognitives et non cognitives des individus (CUNHA et HECKMAN, 2007). Notons que l'on pourrait également considérer les facteurs apportés par les pairs (les caractéristiques des élèves de la classe ou du voisinage) comme une troisième catégorie de ressources (GOUX et MAURIN, 2007). On supposera ici, pour simplifier, que ces ressources rentrent dans l'apport de l'environnement familial au sens large.

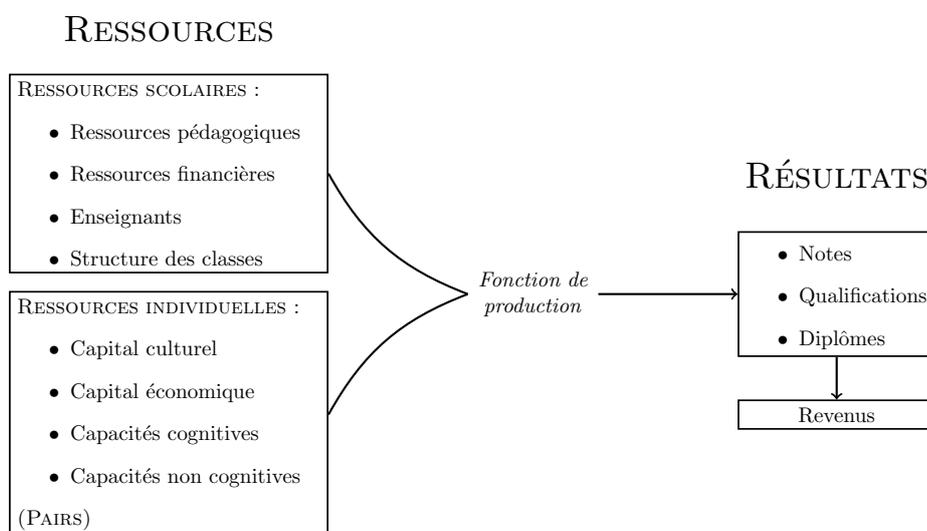


FIGURE 1 – La fonction de production de l'éducation

L'importance relative de ces deux types de ressources dans la fonction de production de l'éducation est sujet à un long débat dans la littérature. À la suite de la démocratisation scolaire, cette question est née du constat que l'égalité d'accès à l'éducation n'impliquait pas nécessairement l'égalité de réussite scolaire. Le rapport « *Equality of Educational Opportunity* », plus connu sous le nom de rapport Coleman (COLEMAN et al., 1966), commandé par le Congrès américain sur le thème de l'inégalité des chances éducatives selon, notamment, l'origine ethnique des élèves, a ainsi montré que les ressources scolaires expliquaient en réalité peu de la variation de la réussite des élèves aux États-Unis, une fois tenu compte du milieu socio-économique². Les différences de milieu socio-économique ainsi que les

²Le rapport montre cependant que les ressources scolaires sont un facteur explicatif plus

caractéristiques des pairs expliquaient une part plus importante des différences de résultats scolaires. Ce résultat a donné lieu à une importante littérature cherchant à évaluer le rôle de ces différentes ressources sur les performances scolaires. Dans un ensemble de méta-analyses largement citées, HANUSHEK (1981, 1986, 1989, 1997, 1998, 2003) affirme qu'il n'existe pas, dans la littérature, de relation systématique entre ressources scolaires et résultats des élèves aux États-Unis. L'auteur en conclut que, pour améliorer les résultats scolaires, on ne peut attendre beaucoup des dépenses d'éducation. Ce résultat a été largement critiqué. En réutilisant le corpus d'études analysé par Hanushek, plusieurs auteurs (DEWEY, HUSTED et KENNY, 2000; GREENWALD, HEDGES et LAINE, 1996; HEDGES, LAINE et GREENWALD, 1994; KRUEGER, 2003) ont contesté la méthodologie adoptée (*vote counting*, absence de règle de décision), l'absence de prise en compte de l'endogénéité des ressources scolaires et la sélection des données et ont montré l'existence d'une relation positive et significative entre les ressources scolaires et les résultats. Ce débat ne s'est pas limité au contexte américain et plusieurs auteurs se sont interrogés sur l'impact des différences de ressources scolaires entre pays (voir par exemple WÖSSMANN, 2005 pour les pays d'Europe de l'Ouest, OECD, 2013 pour les pays de l'OCDE, WÖSSMANN, 2003 pour une revue internationale. Voir aussi VIGNOLES et al., 2000 pour une revue de littérature notamment sur le Royaume-Uni.).

La raison de la longévité de ce débat est sans doute qu'il pose plusieurs questions fondamentales pour l'économie de l'éducation. Du point de vue de la théorie économique, comment expliquer qu'une plus grande quantité de ressources ne se traduise pas dans les résultats individuels? Faut-il reconsidérer la fonction de production de l'éducation? Mais surtout, du point de vue des politiques éducatives, faut-il, comme le suggère Hanushek, considérer qu'agir sur les ressources scolaires revient à « gaspiller de l'argent dans les écoles »³? Faut-il mettre en place des incitations ciblant plus directement les individus?

Quelles politiques éducatives ?

Agir directement sur les ressources mobilisées par les individus et les familles n'est pas chose facile. Depuis la fin des années 1990, de nombreux programmes de transferts monétaires conditionnés à la scolarisation des enfants (*conditional*

important pour les élèves de minorités ethniques que pour les autres élèves.

³« Throwing money at school »

cash transfers) ont vu le jour. L'un des plus connus est le programme mexicain *Oportunidades* (anciennement PROGRESA), mis en place à partir de 1997, qui consiste à fournir une aide financière aux familles pauvres, sous condition que les enfants aillent régulièrement à l'école. Ces transferts augmentent significativement le taux de scolarisation des enfants dans ces ménages (ATTANASIO, MEGHIR et SANTIAGO, 2012 ; SCHULTZ, 2004). Au Royaume-Uni, l'allocation scolaire *Education Maintenance Allowance* cible les jeunes de 16 à 19 ans issus de ménages à faible revenu pour les inciter à poursuivre des études. Là encore, les allocations, de 10 à 30 livres sterling (soit environ 14 à 40 euros) par semaine, conduisent significativement plus de jeunes à poursuivre des études (MIDDLETON et al., 2005). Au delà de l'aspect financier, certains programmes cherchent à inciter les parents à s'investir plus dans l'éducation de leurs enfants. En France, c'est le cas par exemple de la « mallette des parents », dispositif expérimenté dans une quarantaine de collèges de l'académie de Créteil en 2008-2009, puis élargi à d'autres écoles et collèges, notamment en éducation prioritaire. Le programme consiste à convier les parents à des réunions d'information sur le fonctionnement de l'institution scolaire, le déroulement de la scolarité de leurs enfants et sur les façons de les accompagner dans leur scolarité. AVVISATI et al. (2013) montrent que les parents bénéficiant du programme augmentent leur investissement dans l'éducation des enfants, à la fois à l'école et à la maison. L'attitude des enfants à l'école est améliorée, même si leurs résultats scolaires ne sont pas significativement plus élevés.

Parce qu'il est plus aisé de cibler plus globalement des établissements, la plupart des politiques éducatives sont mises en place au niveau des écoles, voire au niveau d'un ensemble d'établissements ou d'un quartier. C'est le cas des politiques d'éducation prioritaire, dont l'objectif est de compenser les inégalités de réussite scolaire entre les élèves issus de milieux sociaux défavorisés et les autres élèves. Elles ciblent traditionnellement des établissements identifiés comme scolairement et socialement désavantagés. Pour cette raison, elles sont parfois qualifiées de dispositifs « zonés ». En France, l'éducation prioritaire cible en réalité des établissements et non pas des zones (comme le terme « zones d'éducation prioritaire » pourrait le laisser penser). Toutefois ces dispositifs recoupent souvent géographiquement les politiques de la ville, qui ciblent plus largement des quartiers. Les politiques éducatives territorialisées rentrent en ce sens dans la dénomination « *place-based people strategies* » proposée par l'économie urbaine (LADD, 1994 ; NEUMARK et SIMPSON, 2015), c'est-à-dire qu'elles ciblent des établissements ou des quartiers, mais

dans l'intention de venir en aide aux populations scolarisées dans ces établissements ou résidant dans ces quartiers.

La question qui se pose alors est : qui bénéficie effectivement des politiques territorialisées ? D'une part, puisque ces dispositifs sont mis en place au niveau des établissements et non directement au niveau des ménages ou des élèves, il n'est pas sûr qu'ils parviennent à atteindre les individus initialement visés. D'autre part, les ménages sont susceptibles de réagir à une modification de l'allocation spatiale des ressources en changeant de localisation ou d'établissement, modifiant ainsi la population effectivement ciblée. En effet, identifier un établissement comme socialement et scolairement défavorisé est susceptible de créer un signal négatif et des stratégies d'évitement. Au contraire, il se pourrait que certains individus se relocalisent pour bénéficier d'un dispositif zoné. Comment alors évaluer les effets de ces dispositifs ?

Comment les évaluer ?

La première difficulté dans l'évaluation des dispositifs éducatifs provient de l'endogénéité des ressources scolaires, c'est-à-dire du fait que des facteurs inobservés par l'économiste sont susceptibles d'affecter à la fois les ressources éducatives et les performances scolaires des élèves, biaisant ainsi l'analyse de la relation entre ces deux grandeurs. Par exemple, les parents peuvent choisir, dans une certaine mesure, les ressources scolaires (*via* le choix d'établissement, les choix d'options ou le choix de résidence). Supposons que les parents les mieux à même de choisir les meilleures ressources pour leurs enfants (par exemple les parents les mieux informés ou les plus éduqués) soient aussi ceux dont les enfants réussissent le mieux à l'école. Alors il existe une relation positive entre ressources scolaires et performance des élèves, même si de meilleures ressources ne conduisent pas les élèves à obtenir de meilleurs résultats. Autrement dit, la relation positive entre les ressources et les résultats scolaires n'est pas causale, elle est due à la corrélation positive entre les caractéristiques des parents et les performances scolaires de leurs enfants. Il est également fréquent que les ressources soient réparties de façon à donner plus à ceux qui ont des difficultés scolaires (en les mettant dans des classes plus petites par exemple). Dans ce cas, la relation est négative ; l'accès à de meilleures ressources est associé à de moins bons résultats scolaires, non pas par causalité, mais par le fait que les élèves en difficulté ont accès à de meilleures ressources. Évaluer la fonction de production de l'éducation

nécessite de tenir compte de ces phénomènes pour isoler l'effet propre des ressources scolaires sur les performances éducatives.

Pour évaluer l'effet causal des dispositifs éducatifs sur la réussite, il faudrait donc pouvoir observer une modification dans les ressources scolaires qui soit complètement indépendante à la fois des choix de (re)distribution des ressources et des choix des individus. La meilleure manière d'observer une telle variation est d'allouer aléatoirement les moyens. Si une partie des élèves bénéficient de certaines ressources de façon aléatoire (par exemple par tirage au sort) et d'autres non, alors la comparaison des résultats scolaires de ces élèves permet de mesurer l'effet propre des ressources en question sur la performance scolaire. Ce type de dispositif est le cadre idéal pour l'évaluation de la fonction de production de l'éducation, mais il est évidemment très rarement observé. Un exemple connu est le projet STAR (*Student-Teacher Achievement Ratio*), conduit de 1985 à 1989 dans l'État du Tennessee, qui consistait à assigner aléatoirement des élèves de maternelle et leurs enseignants à des classes de petite taille (entre 13 et 17 élèves), à des classes de taille standard (entre 22 et 25 élèves), ou à des classes de taille standard avec un enseignant supplémentaire à plein temps. La comparaison de ces trois groupes permet de montrer que les élèves scolarisés dans les petites classes obtiennent de meilleurs résultats que les autres (KRUEGER, 1999). Le bénéfice d'une petite classe s'avère par ailleurs plus important pour les élèves issus de minorité ethnique et pour les élèves issus de milieux socialement désavantagés.

En l'absence de telles expériences randomisées, il est parfois possible d'observer des modifications dans les ressources, qui, bien que n'étant pas aléatoires, peuvent être supposées indépendantes des performances scolaires des élèves. La règle de Maimonides a ainsi permis d'évaluer l'effet de la taille des classes dans les écoles publiques israéliennes (ANGRIST et LAVY, 1999). Cette règle veut que la taille de classe maximale soit de 40 élèves, si bien que l'on ouvre une classe supplémentaire à partir de 41 élèves. Cette règle arbitraire permet de comparer des élèves scolarisés dans des classes de taille variable uniquement du fait de la règle et pas de considérations dépendant des performances des élèves dans les classes. Les auteurs montrent alors qu'une réduction de la taille des classes dans l'enseignement élémentaire conduit à de meilleurs résultats. En appliquant cette même règle au contexte français, PIKETTY et VALDENAIRE (2006) ont également montré que des classes plus petites améliorent les résultats des élèves, surtout dans l'enseignement élémentaire. HOXBY (2000) a proposé une méthode différente consistant à utiliser les

variations de la taille des cohortes d'élèves dans les écoles. En utilisant cette méthode, l'auteur ne trouve pas d'effet significatif de la réduction de la taille des classes dans des écoles élémentaires aux États-Unis. En France, BRESSOUX, KRAMARZ et PROST (2009) utilisent le système d'affectation des jeunes enseignants pour montrer que la formation des professeurs des écoles améliore de façon significative les résultats des élèves en mathématiques et que les classes de petite taille ont un effet positif et significatif sur les résultats.

Lorsque l'on tient compte de l'endogénéité des ressources scolaires, celles-ci semblent donc bien souvent permettre d'améliorer les résultats des élèves. Bien évidemment, les effets sont probablement très différents d'un pays à l'autre, d'une école à l'autre et d'un élève à l'autre. Il semble également que les effets sont souvent plus importants pour les élèves les plus jeunes et pour les élèves les plus défavorisés. Néanmoins, il existe un faisceau d'analyses démontrant les effets positifs d'une amélioration des ressources scolaires.

Le deuxième écueil dans l'évaluation de la fonction de production de l'éducation est l'interdépendance des ressources scolaires et individuelles. Tout d'abord, les individus ne sont pas de simples consommateurs de l'éducation, ils ne se contentent pas de profiter des ressources mises à leur disposition ; ils sont acteurs dans la détermination de ces ressources. Le modèle proposé par TIEBOUT (1956) suggère que les individus, en tant que consommateurs et électeurs, participent à la provision de biens publics. D'après ce modèle, les individus choisissent le lieu de résidence qui correspond le mieux à leurs préférences en termes de biens publics, notamment selon la qualité des écoles (EPPLÉ et ROMANO, 2003), et votent pour déterminer la quantité des dépenses permettant de financer ces biens publics au niveau local. Les individus sont donc attentifs à l'offre scolaire locale dans leur choix de résidence. Plusieurs études montrent ainsi que les ménages sont prêts à payer plus cher pour de meilleures écoles et que ceci se traduit par des prix de l'immobilier plus élevés (voir notamment BLACK et MACHIN, 2011, pour une revue de cette littérature). Par exemple, l'étude de BLACK (1999) montre que, au milieu des années 1990 dans la banlieue de Boston, les ménages étaient prêts à payer leur logement 2,5 % plus cher pour avoir accès à des écoles obtenant des résultats en moyenne 5 % plus élevés. Concernant la France, FACK et GRENET (2010) montrent qu'une augmentation d'un écart-type de la moyenne au brevet des collèges (soit 1,44 point sur 20) correspond à une augmentation de 40 % des prix de l'immobilier dans le secteur d'un collège à Paris. Les auteurs trouvent par ailleurs que cet effet est d'autant plus faible que

l'offre de collèges privés dans le secteur est importante.

Puisque les individus sont acteurs dans la détermination des ressources scolaires, alors une modification de ces ressources est susceptible de modifier les comportements individuels en matière d'éducation. Cette question a été abordée seulement assez récemment dans la littérature, mais plusieurs études ont montré que les parents modifient leur investissement, à la fois personnel et financier, en réponse à une modification des ressources scolaires. En étudiant un programme scolaire national en Zambie et une expérience randomisée en Inde, DAS et al. (2013) montrent par exemple que les ménages réduisent leurs dépenses d'éducation lorsqu'ils anticipent que leur école va recevoir une subvention. Par ailleurs, HOUTENVILLE et CONWAY (2008) montrent que l'investissement des parents dans l'éducation des enfants (mesuré par la fréquence des discussions sur le thème de l'école ou la participation à des réunions scolaires) est corrélé négativement aux ressources scolaires, suggérant de possibles effets d'éviction, les parents étant susceptibles de réduire leurs efforts en réponse à une augmentation des ressources. Dans le contexte du système scolaire secondaire en Roumanie, POP-ELECHES et URQUIOLA (2013) montrent également que les parents aident moins leurs enfants à faire leurs devoirs lorsque ceux-ci sont scolarisés dans un meilleur lycée.

En outre, les individus sont susceptibles de réagir à de l'information sur les résultats de la fonction de production. FIGLIO et LUCAS (2004) ont ainsi cherché à savoir si, à qualité des écoles fixée, donner une information supplémentaire aux familles sur les écoles était en mesure de modifier leurs comportements. Pour répondre à cette question, les auteurs analysent l'effet d'un dispositif mis en place en 1999 dans l'État de Floride consistant à évaluer chaque année toutes les écoles de l'État sur la base de notes allant de A à F. En étudiant le marché de l'immobilier dans les années suivant la mise en place de ce système de notation, les auteurs trouvent que, à qualité donnée, recevoir un A plutôt qu'un B est associé à une augmentation de 19,5 % des prix des ventes dans le secteur de l'école. Ils montrent par ailleurs que, à court terme, les meilleurs élèves déménagent pour aller dans les écoles obtenant les meilleures notes. HASTINGS et WEINSTEIN (2008) utilisent le contexte du *No Child Left Behind Act*, qui autorise à changer d'école les familles dont l'école de secteur ne voit pas les résultats des ses élèves progresser pendant deux années consécutives. Les auteurs montrent que le fait de donner de l'information aux parents concernés sur les résultats des autres écoles avoisinantes augmente la proportion de familles qui choisissent une école obtenant de meilleurs résultats.

Est-ce que ces constats doivent dissuader les décideurs publics de mettre en place des politiques éducatives? Faut-il éviter d'interférer avec les décisions individuelles de scolarisation? Comment tenir compte de ces réactions individuelles dans l'évaluation de la fonction de production de l'éducation? L'évaluation des politiques éducatives est doublement complexe. Non seulement il est nécessaire de tenir compte de l'endogénéité des ressources présentée plus haut, mais il faut également prendre en compte le fait que la modification des ressources en elle-même est susceptible d'affecter les comportements des individus. Dans cette thèse, nous proposons de dépasser le débat sur l'importance relative des ressources scolaires et de l'environnement familial pour considérer l'interdépendance de ces deux types de ressources. Plutôt que de savoir qui, de l'école ou de la famille, contribue le plus à la fonction de production de l'éducation, nous proposons de nous interroger sur la façon dont ces deux acteurs interagissent pour produire le capital humain.

Contribution de la thèse

L'objectif de cette thèse est de tenir compte des réactions individuelles pour évaluer la fonction de production de l'éducation. Pour cela, nous proposons trois analyses, qui s'appuient à la fois sur des méthodes économétriques permettant d'isoler des effets causaux et sur des données au niveau individuel. Nous verrons, dans le chapitre 1, que les dispositifs d'éducation prioritaire en France conduisent les familles à modifier leurs choix d'établissements. Nous verrons également que cette modification des comportements n'est pas indépendante des caractéristiques socio-économiques des familles, ce qui amène à sous-estimer les résultats des politiques d'éducation prioritaire si l'on ne tient pas compte des modifications dans le choix d'établissement. Dans le chapitre 2, nous verrons qu'une modification de l'offre scolaire locale modifie les décisions d'orientation des individus. Enfin, dans le chapitre 3, nous verrons que la connaissance de leurs notes à l'examen est susceptible de modifier la façon dont les étudiants évaluent les qualités pédagogiques de leurs enseignants.

Méthodes et sources utilisées

Les données utilisées dans cette thèse, que nous présenterons plus en détail dans les chapitres suivants, doivent permettre une analyse précise des comportements individuels, des résultats et des ressources scolaires. Il s'agit donc de données individuelles décrivant, au niveau le plus fin possible, les choix scolaires et les

résultats des élèves. Dans les chapitres 1 et 2, les données sont issues de fichiers administratifs de la Direction de l'évaluation, de la prospective et de la performance (DEPP) du Ministère de l'éducation nationale. Les fichiers anonymisés d'élèves pour la recherche et les études (chapitre 1) et les bases SCOLARITÉ (chapitre 2) recensent chaque année l'ensemble des élèves scolarisés dans un établissement de l'enseignement secondaire dépendant du Ministère de l'éducation nationale en France. L'intérêt de ces données exhaustives est donc d'offrir une vision globale du système scolaire français. Nous utilisons par ailleurs la base centrale des établissements (BCE) contenant des informations précises sur la totalité des établissements scolaires français, et notamment l'adresse postale de chaque établissement. Sur la base de ces adresses, chaque établissement a été géolocalisé. Il est alors possible non seulement de localiser précisément les établissements de scolarisation des élèves, mais également de reconstruire l'offre scolaire locale de chacun des élèves présents dans les bases individuelles⁴. Ces données géolocalisées permettent donc de connaître l'allocation spatiale des ressources scolaires par rapport aux élèves.

Les données utilisées dans le chapitre 3 proviennent d'une grande école française. Ces données sont quasiment uniques en France (pour une application différente sur une autre institution parisienne, voir le travail de BORING, 2015) ; elles permettent d'observer, pour chaque année scolaire et pour chaque cours proposé, les évaluations anonymes et individuelles remplies par les étudiants. Elles recensent également séparément l'ensemble des notes obtenues par les étudiants pour chacun des cours. Bien qu'étant anonymes et ne permettant donc pas une analyse au niveau individuel, ces données fournissent un outil rare pour analyser la façon dont les étudiants évaluent les qualités de l'enseignement qu'ils reçoivent. Elles permettent notamment d'observer la date précise à laquelle chaque étudiant évalue chaque cours.

Pour estimer l'effet causal d'une modification des facteurs de la fonction de production de l'éducation, nous utilisons, dans cette thèse, plusieurs outils économétriques. Une allocation aléatoire des ressources étant évidemment très rarement

⁴Dans une version précédente de l'analyse présentée au chapitre 1, nous avons utilisé les panels d'élèves de la DEPP. L'intérêt de ces panels est de coupler aux informations administratives des données très riches issues de questionnaires individuels. Mais l'immense avantage des panels d'élèves est que nous avons pu géolocaliser l'adresse de résidence des élèves, ce qui n'est pas possible dans les données exhaustives, l'adresse des élèves n'étant pas connue. Cependant, les méthodes utilisées dans le chapitre 1 réclamant une grande puissance statistique, nous avons dû travailler finalement sur les données exhaustives pour augmenter considérablement la taille des échantillons et la puissance des estimations.

observée, il faut mettre en œuvre des stratégies permettant de tenir compte à la fois de l'endogénéité des ressources et des réactions individuelles.

Régressions sur discontinuité

Le premier chapitre utilise des régressions sur discontinuités pour estimer l'effet d'un dispositif d'éducation prioritaire sur les choix d'établissement et la réussite scolaire. La stratégie d'estimation repose sur les règles d'éligibilité au dispositif. La localisation des établissements de l'éducation prioritaire n'est évidemment pas aléatoire ; ceux-ci sont, par définition, situés dans des voisinages scolairement et socialement défavorisés. La sélection des établissements dans le dispositif étudié ici s'est ainsi faite sur la base de critères académiques et sociaux évalués au niveau national. Pour comparer des élèves comparables, nous utilisons alors les seuils d'éligibilité au dispositif dans l'établissement le plus proche. Plus précisément, nous utilisons le fait que les élèves résidant à proximité d'un établissement situé juste au-dessous des seuils d'éligibilité et les élèves résidant à proximité d'un établissement situé juste au-dessus des seuils sont similaires, mais les derniers ont une probabilité plus élevée que les premiers d'être scolarisés en éducation prioritaire.

Différences de différences

Le deuxième chapitre met en œuvre des différences de différences généralisées pour analyser l'effet de l'ouverture d'un nouveau lycée sur les décisions d'orientation des élèves. Il s'agit cette fois d'utiliser la variation dans le temps et dans l'espace des ouvertures d'établissements pour tenir compte du fait que les lycées n'ouvrent pas n'importe où sur le territoire. Le principe est de comparer, d'une part, les choix d'orientation des élèves d'un même collège avant et après l'apparition d'un nouveau lycée et, d'autre part, les choix des élèves à une date donnée dans deux collèges différents, dont l'un voit apparaître un nouveau lycée et l'autre pas. Cette double différence permet de mesurer l'effet d'un changement de l'offre scolaire tout en contrôlant à la fois les caractéristiques propres au collège d'origine et qui pourraient expliquer les choix scolaires, mais aussi les tendances temporelles dans les choix scolaires qui sont communes à tous les collèges.

Effets fixes et variables instrumentales

Le troisième chapitre utilise également des données longitudinales pour étudier l'effet sur l'évaluation des enseignants par les étudiants d'une information sur la qualité de l'enseignement. L'observation de cohortes répétées d'étudiants pour chaque cours nous permet d'utiliser des effets fixes enseignants. Le principe est de contrôler les caractéristiques inobservées des enseignants (le charisme, la motivation), qui affectent à la fois les résultats des étudiants et leur évaluation. Nous proposons également des variables instrumentales pour tenir compte de l'endogénéité de la note obtenue à l'examen sur l'évaluation du cours. Nous utilisons enfin un modèle semi-paramétrique pour analyser l'évolution des évaluations dans le temps et selon l'information dont les étudiants disposent.

Quelle efficacité des politiques d'éducation prioritaire ?

Le premier chapitre s'interroge sur l'efficacité des dispositifs d'éducation prioritaire. En France, les politiques d'éducation prioritaire se montrent relativement impuissantes à améliorer les résultats scolaires des élèves (BEFFY et DAVEZIES, 2013 ; BÉNABOU, KRAMARZ et PROST, 2004, 2009 ; CAILLE, 2001 ; MEURET, 1994 ; MOISAN et SIMON, 1997). Un argument susceptible d'expliquer ces mauvais résultats est qu'en identifiant des établissements comme socialement et scolairement défavorisés, les dispositifs d'éducation prioritaire donnent un signal négatif et créent des stratégies d'évitement des établissements sélectionnés (cf. figure 2). Dans la première partie de ce chapitre, issue d'une étude co-écrite avec Laurent DAVEZIES, nous cherchons à savoir si cet effet existe dans le cas du dispositif « réseaux ambition réussite » (RAR), qui a ciblé les collèges les plus défavorisés de l'éducation prioritaire entre 2006 et 2011. Notre question est double. Le programme RAR a-t-il induit des effets d'évitement scolaires ? Ces effets sont-ils susceptibles de créer ou de renforcer la ségrégation sociale entre collèges ?

En utilisant une base de données exhaustive des élèves entrés en sixième au moment de l'entrée en vigueur du dispositif, nous mettons en œuvre une estimation par régression sur discontinuité et nous trouvons qu'habiter à proximité d'un collège RAR diminue la probabilité d'être scolarisé dans le collège le plus proche et augmente la probabilité d'être scolarisé dans un établissement privé. Nous trouvons également que le dispositif RAR renforce la ségrégation sociale entre collèges, au sens où les élèves issus de milieux plus favorisés ont une probabilité plus élevée que les autres de

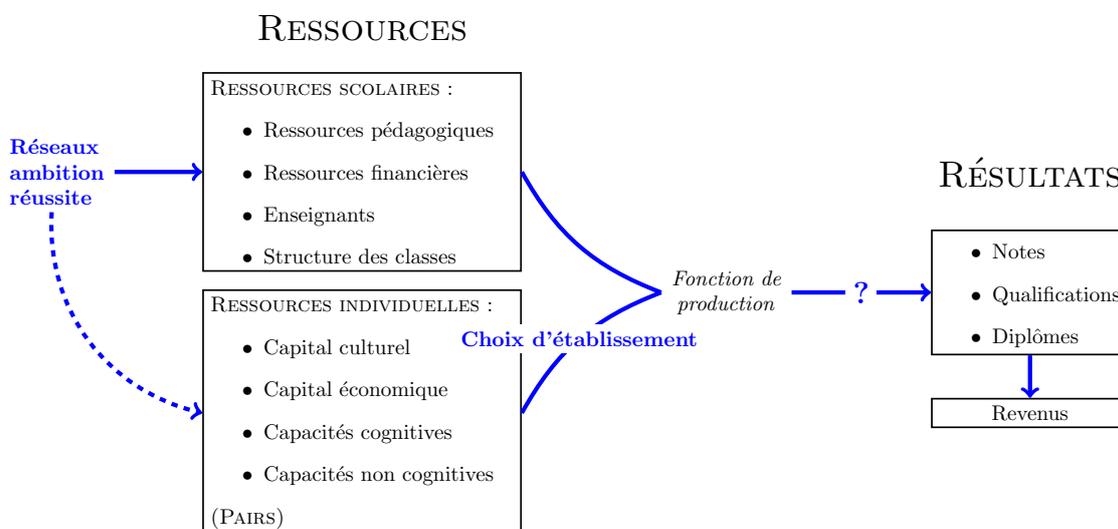


FIGURE 2 – Schéma chapitre 1

contourner les établissements RAR lorsqu'ils habitent à proximité d'un collège RAR. Ceci contribuerait à expliquer pourquoi les dispositifs zonés d'éducation prioritaire n'ont pas les effets attendus, puisque les élèves issus des familles socialement plus favorisés (qui ont de plus grandes chances de réussir à l'école en moyenne) contournent les établissements ciblés.

Dans une deuxième partie, issue d'une analyse co-écrite avec Jean-Paul CAILLE et Laurent DAVEZIES, nous cherchons à évaluer l'effet du dispositif RAR sur les résultats scolaires des élèves, lorsque les stratégies d'évitement sont prises en compte. En étudiant plusieurs dimensions de réussite scolaire (résultats au brevet des collèges, orientation en fin de collège), nous ne trouvons pas d'effet de la scolarisation en RAR sur les résultats, ni sur les trajectoires scolaires des élèves. Cette absence d'effet global pourrait cependant être due à des effets différenciés sur certaines sous-populations d'élèves. Ainsi, dans certains collèges, les filles semblent bénéficier du dispositif, à l'inverse des garçons et des élèves issus de catégories sociales favorisées.

Comment l'offre locale affecte les choix d'orientation ?

Les politiques éducatives de ces dernières années tendent à donner plus de liberté aux familles dans le choix des ressources éducatives mobilisées pour leurs enfants. Un certain nombre de réformes, comme l'assouplissement de la carte scolaire en

France, vont dans ce sens. Ces réformes sont censées conduire à une amélioration du système scolaire à la fois en augmentant la compétition entre les établissements et en créant de meilleurs appariements entre les établissements et les élèves. L'efficacité de ces réformes reposent toutefois sur la double hypothèse i) que les individus sont contraints dans leurs choix et ii) qu'ils sont les mieux à même de choisir les meilleures ressources. Nous proposons de tester ces hypothèses dans le cadre des choix d'orientation qui ont lieu à la fin du collège en France.

L'objectif du deuxième chapitre, issu d'une étude co-réalisée avec Meryam ZAIEM et Son Thierry LY, est de déterminer dans quelle mesure les décisions individuelles d'orientation varient en fonction de l'offre de formation (voir figure 3). Nous nous intéressons plus précisément aux choix d'orientation des élèves à la fin du collège et nous étudions l'effet de l'offre scolaire au voisinage du collège des élèves. Notre question est double. Est-ce que les élèves sont contraints dans leurs choix par l'offre scolaire locale ? Est-ce que leur orientation change lorsque l'offre scolaire est modifiée par la création d'un nouvel établissement. Pour répondre à ces questions, nous utilisons des données exhaustives en coupes répétées sur les cohortes d'élèves en classe de troisième de l'année scolaire 2002-2003 à l'année scolaire 2011-2012 et nous exploitons les variations dans le temps et dans l'espace des ouvertures de lycées au voisinage de leurs collèges.

En utilisant une méthode de différences de différences, nous trouvons que l'ouverture d'un nouveau lycée public dans le voisinage d'un collège augmente la proportion d'élèves qui continuent en second cycle du secondaire. Nous en concluons que l'offre scolaire exerce effectivement une contrainte sur les décisions individuelles d'orientation. De plus, nous trouvons que l'effet est porté par les élèves qui s'orientent en voie professionnelle.

Comment évaluer la qualité des enseignements ?

L'évaluation des enseignements par les étudiants est de plus en plus utilisée dans l'enseignement supérieur en France. Pourtant, il existe peu d'études sur la façon dont les étudiants forment ces évaluations. L'objectif du troisième chapitre, issu d'une étude co-réalisée avec Ronan LE SAOUT, est d'analyser la relation entre la note à l'examen et l'évaluation des enseignements par les étudiants (voir schéma figure 4). Comment les étudiants réagissent à une information sur la qualité de l'enseignement, *via* leur note ? L'analyse de la relation entre les résultats des étudiants et l'évaluation d'un enseignement est rendue difficile par le fait que des

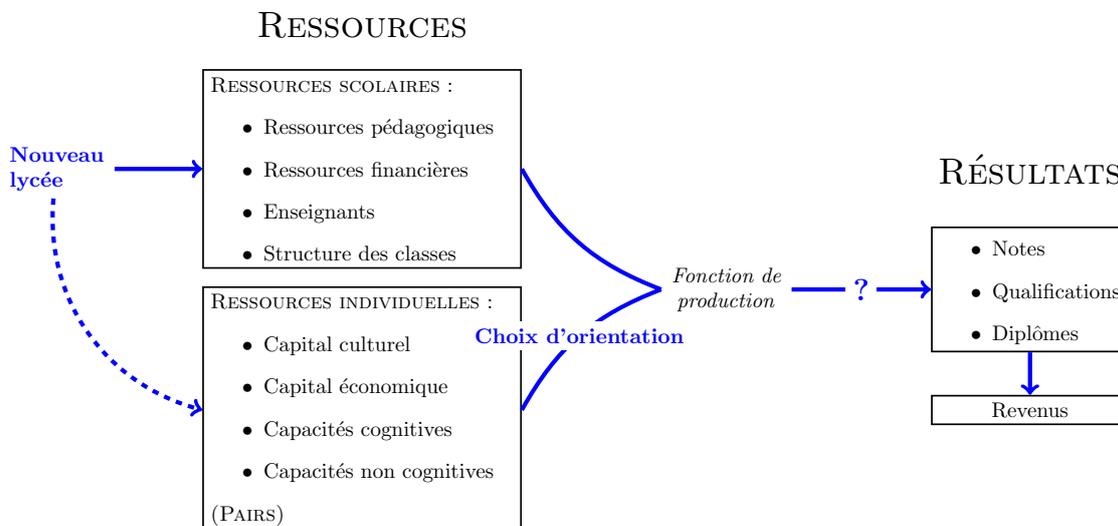


FIGURE 3 – Schéma chapitre 2

caractéristiques inobservées des enseignants (comme le charisme ou la motivation) et des étudiants (comme la facilité à apprendre ou l'intérêt pour la matière) déterminent à la fois les notes obtenues et les évaluations. En outre, les notes et les évaluations sont vraisemblablement déterminées simultanément parce qu'un bon enseignement devrait à la fois conduire les étudiants à obtenir de bonnes notes et à donner de bonnes évaluations.

En exploitant une base de données originale des résultats d'une grande école française sur 8 années scolaires et près de 100 matières, nous utilisons des effets fixes professeurs et des régressions par variables instrumentales pour identifier l'objectivité des étudiants lorsqu'ils évaluent les qualités pédagogiques des enseignants. Nous utilisons également la variabilité de la date de l'évaluation par les étudiants. Nous montrons qu'une notation moins sévère se traduit par une meilleure évaluation moyenne de la pédagogie de l'enseignant. Les effets sont néanmoins quantitativement faibles. Nous trouvons également que les étudiants tiennent compte de l'information contemporaine dont ils disposent au moment où ils évaluent ; ils donnent de moins bonnes évaluations après avoir passé l'examen final et de meilleures évaluations après avoir obtenu leur note.

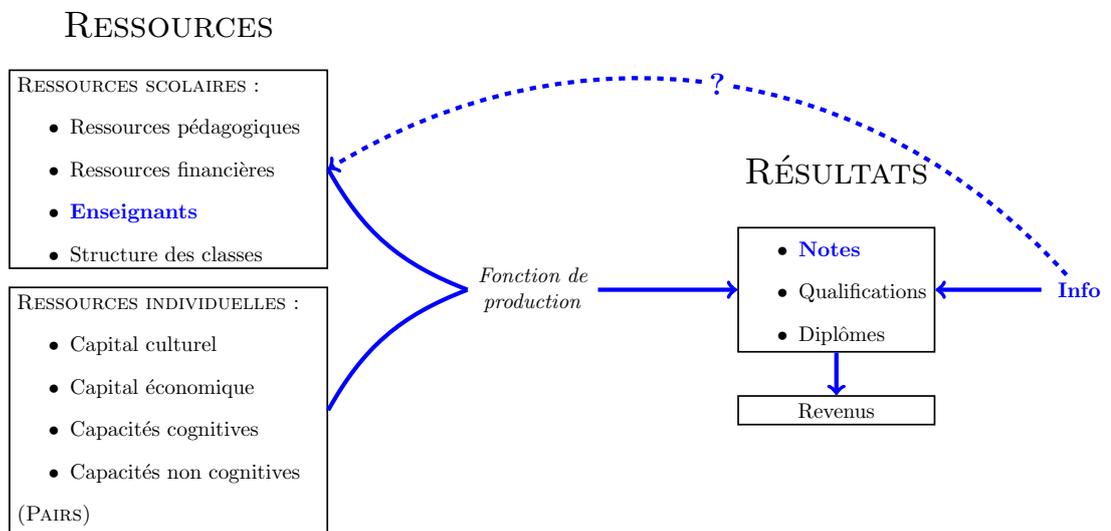


FIGURE 4 – Schéma chapitre 3

Chapter 1

More harm than good? Sorting effects in a compensatory education program

This chapter is based on joint works with Jean-Paul CAILLE and Laurent DAVEZIES¹

Introduction

Compensatory education policies aim at offsetting educational inequalities between socially and academically disadvantaged children and more advantaged ones. These policies first appeared in the 1960's, when mass schooling and equal access to education were found to be insufficient to ensure equal opportunity in developed countries. The fundamental idea is to provide some sub-population with additional resources in order to achieve equal opportunity through unequal treatment. Such education programs traditionally target schools in socially and academically deprived areas. For that reason, they are referred to as place-based (or school-based) programs, as opposed to individual-based ones. Title I of the Primary and Secondary Education Act in the US, Education Priority Areas in the UK, or "*Zone d'éducation prioritaire*" (ZEP) in France are examples, among many, of such policies. These programs exist worldwide, and they usually represent a significant part of public spending in education. The sole Title I program, for instance, represents around

¹Caille, Davezies, and Garrouste (2015) and Davezies and Garrouste (2014)

14 billion dollars per fiscal year. In France, compensatory education corresponds to about 1 billion euros each year. In both cases, the additional spending over the number of recipients represents about 10% of the annual spending per pupil.

Providing underprivileged schools with (sufficiently) more resources is expected to improve pupils' performance and, ultimately, to close the educational gap.² The empirical evidence, however, is mixed. First, the empirical relationship between school resources and pupils' performance is the subject of an extensive debate. Hanushek (2003) finds it to be weak, whereas Krueger (2003), Hedges, Laine, and Greenwald (1994), and Dewey, Husted, and Kenny (2000) find a positive and quite large link. Second, empirical evaluations of school-based compensatory education measures have not reached a consensus, either. Some programs, notably when they intervene very early in primary or in pre-primary education, have both positive and somewhat large effects on pupils' performance (Dobbie and Fryer, 2011; Shapiro and Moreno Treviño, 2004; Tokman, 2002). But other place-based compensatory education programs have limited results. Title I in the US, for example, proved to have a modest overall impact on pupils' outcomes (Borman and D'Agostino, 1996) and may even have had adverse effects, according to Klaauw (2008). In the UK, Machin, McNally, and Meghir (2004) and Machin, McNally, and Meghir (2010) provide an evaluation of the Excellence in Cities program and find a positive but modest impact on pupils' performance (for math scores only) and a reduction of absenteeism. In France, Bénabou, Kramarz, and Prost (2009) show that the ZEP program had no impact on pupils' achievement. Beffy and Davezies (2013) find a negative impact of the "*Réseaux ambition réussite*" program on mean school academic achievement.

Some factors may explain why such policies have limited empirical results. First, some claim that financial support is scattered across too many schools for an inadequately low final amount per pupil. Van der Klaauw (2008) explains, for instance, that "almost all schools are eligible to at least some Title I funds." In France in 2013, more than 1 pupil in 7 was enrolled in a compensatory education school. Second, if spending per pupil were sufficient, it could be misallocated. Since these programs target schools and not individuals directly, they may fail to reach

²Compensatory education policies were originally not meant to last. In France, for instance, additional resources provided to ZEP schools were supposed to end once the achievement gap closed: "If a sustained action over several years is needed, it would not be advisable to consider permanent assistance [...]" (Minister of National Education, Circular No. 81-536 clarifying the objectives of "*zones prioritaires*," December 1981, authors' translation).

the pupils to whom the benefits would be greatest. Both arguments could explain small estimated effects. However, it seems much more difficult to explain why some empirical studies find negative results.

Evaluating compensatory education programs is a complex task since two effects are likely to bias the analysis. First, by definition, such programs target disadvantaged populations. The selection into the program is often made at the school level, on the basis of social and academic criteria (location in deprived areas, bad academic achievement, large proportion of pupils from ethnic minorities, or from disadvantaged social backgrounds, etc.). Consequently, a naive comparison of treated (those who benefit from the program) and non treated individuals is intrinsically false, since the former were selected because they had poorer academic achievement than the latter on average. In other words, treated pupils perform worse than non treated ones, not because of the program, but because the treated are different *ex ante* from the rest of the population. This effect is commonly referred to as a selection bias. To answer this classical issue, one need to compare the treated group with a control group, i.e. individuals who do not benefit from the program but are similar to the treated ones. This is the purpose of using quasi-experimental settings, instrumental variables (Bénabou, Kramarz, and Prost, 2009), or regressions on discontinuity (Befy and Davezies, 2013; Klaauw, 2008). A second source of bias comes from the fact that compensatory education programs usually target schools and not individuals directly. Because individuals can choose which school to attend, it is difficult to estimate the individual effect of the program on pupils, because one need to account for school and location choices. Individuals may select themselves into (or out of) the program by choosing (or avoiding) a school that benefit from it (Fack and Grenet, 2013; Murat and Thauvel-Richard, 2013). We will refer to this second source bias as a sorting effect. So far, this second type of bias has not been properly accounted for in the litterature.

One main limit of the literature on compensatory education is that it usually focuses on schools and ignores individual adjustments to school-based policies. Yet endogenous sorting of individuals across schools or across neighborhoods is expected to significantly modify place-based policies impacts. A growing theoretical and empirical literature shows first that individuals do adjust to a change in public good provision by moving or changing schools (Ferreira, 2007; Hsieh and Urquiola, 2006; Urquiola, 2005; Urquiola and Verhoogen, 2009), and second, that they incorporate these adjustments in their decisions (Epple, Romer, and Sieg, 2001). In this chapter,

we want to explore the idea that individual responses to school-based policies may mitigate their expected impacts (Nechyba, 2003; Pop-Eleches and Urquiola, 2013) and lead to adverse effects.

More precisely, we analyze the impact of place-based compensatory education on individual sortings across schools. To our knowledge, this is the first attempt to evaluate sorting effects due to a compensatory education program and, in so doing, include equilibrium effects in the analysis of compensatory education policies. Studying such effects is crucial first because, if they are sizeable, then one needs to account for them when evaluating compensatory education policies. Sorting effects would affect the composition of treated schools and thus the estimated impact of the treatment. Targeting schools in academically and socially deprived areas may create a negative signal, resulting in a sorting effect. In that case, more advantaged individuals would tend to avoid selected schools labeled as low-achieving; this could partly explain why estimated effects are small or even negative. The fact that more advantaged pupils select themselves out of treated schools, for instance, may drive part of the overall limited effect. Second, there are non-negligible education policy implications. Policy makers may want to be aware of such equilibrium effects when designing compensatory education programs.

To answer the question, we use the context of the French "*Réseaux ambition réussite*" (RAR) program, which targeted very low-achieving and socially disadvantaged junior high schools between 2006 and 2011. In a first part, we try and assess the causal impact of the program on families' school choice by using an original geocoded individual data set and a regression discontinuity identification strategy. More precisely, we ask whether living in the vicinity of a RAR school affects individual school choice, using an exogenous eligibility scheme of schools into treatment. We find that living near a RAR junior high school decreases the probability to attend the closest school by about 30 percentage points, and symmetrically increases the probability to go to a private school, for pupils living near a school just above the eligibility thresholds. We also find that the RAR program increases social segregation across schools, since pupils from more advantaged backgrounds are more likely to attend a private school when they live near a RAR junior high school. In a second part, we evaluate the causal impact of the program on pupils' academic achievement, once accounted for individual sorting effects. We find no overall effect of the program on pupils' scores at the "*Brevet des collèges*" national exam, nor on their schooling situations five years after entering 6th grade. The absence of average effect may

be due to heterogeneous effects on different sub-populations; we find that, in some schools, being enrolled in a RAR increases (respectively decreases) the probability to pass the "*Brevet*" for girls (respectively for boys). Enrollment in a RAR school may also have negative effects for some pupils from advantaged social backgrounds.

The paper is organized as follows. We briefly describe the French education system and the RAR program in the next section. Section 1.3 presents the data. Section 1.4 presents the effect on individual school choice of the RAR program. Then section 1.5 analyzes the effect on pupils' academic achievement. Section 1.6 concludes and discusses the implications of our results for the evaluation of place-based education policies.

1.1 A brief description of the French education system

In France, education is compulsory for children aged 6 to 16. The French school system consists of 5 years of primary school (ages 6 to 10), then 4 years of junior high school, called "*collège*" (ages 11 to 14), and 3 years of high school, called "*lycée*" (ages 15 to 17).

French primary and secondary education is based on a catchment area system; each pupil is assigned to a public school according to where she lives. Junior high school catchment areas are delimited at the local level by the "*département*" ("*conseil général*"), and each area contains only one junior high school. The catchment area school represents families' default school options. Families also have two outside options: they can either send their child to a private school, or they can ask for another public school through a special dispensation. Most private schools are largely subsidized by the state and follow the same curriculum as public schools (except for religious instruction),³ so they constitute a commonly used outside option. Dispensations, on the other hand, are granted by the regional education authority director on the basis of (in order of priority) medical reasons, scholarship, siblings, distance, and special academic tracks. Pupils living in the catchment area have priority, and dispensations are only accorded if all places were not fulfilled by them.⁴

³Most private schools in France are Catholic schools.

⁴The catchment area system was partly relaxed in 2007. This was supposed to give families more freedom in school choice. The increasing number of dispensation demands resulted in a decrease in the size of RAR schools (Fack and Grenet, 2013). However, as long as there is no more dispensation demands above or below the eligibility thresholds (see below) in the absence of the RAR program, this does not constitute a confounding factor for our analysis.

The 2006 French compensatory education reform created a new structure called the "ambition success" networks ("*Réseaux Ambition Réussite*" or RAR). The program targeted the most disadvantaged junior high schools. Each network consisted of one junior high school, and in the primary and infant schools of the catchment area. The network was managed by an executive committee, composed of the head of the junior high school, the heads of the elementary and infant schools, and the Ministry of National Education inspector responsible for the schools district. The aim of these networks was to build or reinforce the relationships between teachers within the network in order to tackle underachievement in these schools. To achieve this goal, each network had to define an educational project through a four- to five-year contract. Each project had to be built up to reinforce individualized support, develop partnerships with cultural or sports organizations, and strengthen relationships with parents. To reach these objectives, RAR junior high schools were provided with additional resources in order to finance 1 000 extra teachers and 3 000 teaching assistants. The extra teachers, whether primary or secondary, were recruited on the basis of individual RAR projects. Their classroom teaching hours were limited to a halftime service so that they could organize tutoring groups and individual homework assistance, manage teaching assistants, and supervise the relationships with parents. Teaching assistants were in charge of helping pupils both inside and outside the classroom. These additional resources were supposed to enable schools to reduce class size. See Beffy and Davezies (2013) for a detailed evaluation of the program on school resources.

The selection of junior high schools in RAR was made on the basis of objective criteria, which were evaluated at the national level during the 2004/2005 school year. These criteria were the proportion of children from disadvantaged social backgrounds in the school and the proportion of pupils who, upon entering 6th grade, had repeated two grades. More precisely, junior high schools had to have at least 67% of pupils from disadvantaged backgrounds⁵ as well as at least 10% of 6th-grade-level pupils having repeated twice or more, to become eligible to the program.⁶ These thresholds were arbitrarily defined so that selected schools would represent the 5% most socially and academically disadvantaged pupils. The list of eligible schools was approved by

⁵Pupils from disadvantaged backgrounds were defined as children of blue-collar workers, of retired blue-collar workers, of retired white-collar workers, or of the unemployed.

⁶Some regional education authorities used an additional measure of pupils' achievement at the beginning of 6th grade. But since this measure was not available in every school, it is not used in this analysis.

the Minister of National Education. Then some further adjustments were made, and some schools that were eligible did not enter the treatment, whereas some schools below the thresholds did enter it.

The additional cost of RARs has been estimated at nearly 325 million euros for the budget year 2008. It corresponds to about 811 additional euros per pupil, approximately 10% of the annual spending per junior high school pupil. 90% of the extra cost corresponds to the funding of teachers and assistants.

In total, 249 public junior high schools entered the RAR program from the beginning of the 2006 school year. Four additional public schools entered in 2007, as well as 11 private schools.⁷ These schools were located all over the country, mainly in urban areas. Figure 1.1 shows the repartition of public junior high schools in mainland France.⁸

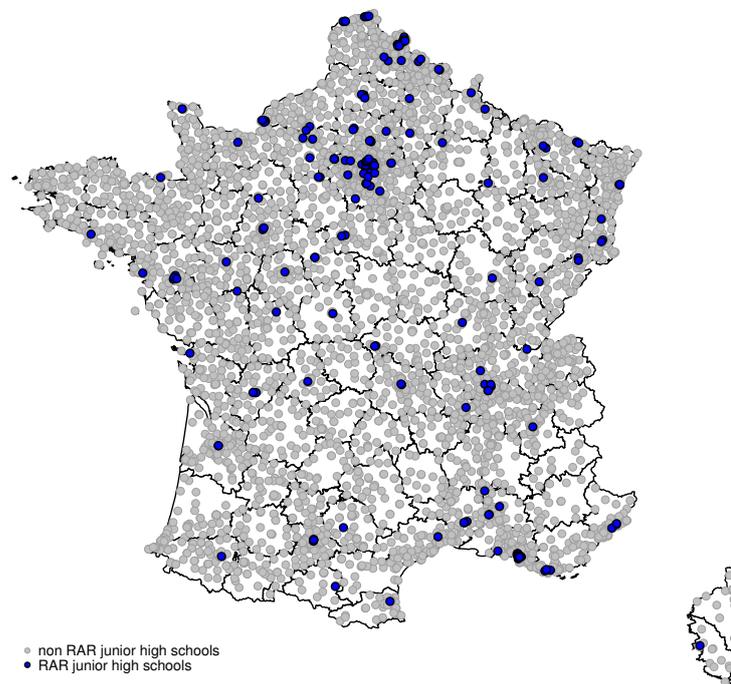
The main objective of the program was to significantly increase the supervising staff in a small number of treated schools. If uniformly distributed among the 249 schools, additional positions would represent about 4 extra teachers, and about 12 extra teaching assistants per school. But because there was no specific follow-up of these schools after the beginning of the program, the actual increase in the number of teachers and teaching assistants need to be estimated. In a preceding study, Befy and Davezies (2013) show that the increase in the number of teaching hours per pupil, and the decrease in class size were not very significant and were less than expected if resources had been uniformly distributed across schools. The authors also find that the proportion of teachers having the highest secondary school teaching certification (*agrégation*) decreased, and that the proportion of teachers more than 55 years old increased.

The expected impact of the program on neighboring families school choices is not clear. On one hand, treated schools may become more attractive, because they benefit from additional resources. Some families may thus try and enroll their child(ren) in these particular schools. On the other hand, the program may signal treated schools as low-achieving and socially disadvantaged. In this case, some families may try and avoid such schools. In the French context, the first scenario is unfortunately not the more credible. Although no one has studied individual

⁷These very few RAR private schools are excluded from the analysis.

⁸Overseas "*départements*" are excluded from the analysis for two reasons. First, the proportion of RAR schools is much higher in overseas "*départements*" than in mainland France, meaning that the eligibility criteria may poorly capture whether they entered the program or not. Second, schools in overseas "*départements*" are often badly geocoded.

Figure 1.1 – Map of public junior high schools in the sample



Source: MEN-MESR DEPP, FAERE 2006 and 2007

school choice in this context before, there is evidence that some families tend to avoid compensatory education schools (Bénabou, Kramarz, and Prost, 2004). One of the objectives of the RAR program was thus to break this negative signalling effect by creating an ambitious and successful environment in treated schools. This first translated into the name of the program. Second, trying to prevent families from avoiding treated schools was mentioned in many RAR schools contracts.

1.2 Theoretical mechanisms

Following Friesen et al. (2012) and Moretti (2011), let us define a simple Bayesian learning model in which each family i chooses a junior high school on the basis of school expected quality. The quality of schools is unobserved, but families hold prior beliefs on school quality, based on observed characteristics. Let us note q_j school j 's unobserved quality. Family i 's prior belief on school j 's quality, conditional on observed characteristics X_j , $C_i(q_j|X_j)$, is assumed to be normally distributed: $C_i(q_j|X_j) \sim \mathcal{N}(X_j'\beta_i, \sigma_{qi}^2)$. The precision $\frac{1}{\sigma_{qi}^2}$ is heterogenous in the population, meaning that some families (for instance the most socially advantaged ones) may get more information on school quality than others. Each family chooses the school j that maximises the expected quality, conditional on observed characteristics. Thus, family i chooses school j if and only if:

$$X_j'\beta_i > X_k'\beta_i, \forall k \neq j$$

Let us now assume that families receive a dichotomous signal $S_j = \mathbf{1}\{S_j^* < 0\}$ on school j , where $S_j = 1$ if school j is RAR, and 0 otherwise. The RAR status is common knowledge, but families do not perfectly observe the selection process of schools into the RAR program. They assume that $S_{ij}^* = q_j + \nu_{ij}$, with $\nu_{ij} \sim \mathcal{N}(0, \sigma_{\nu i}^2)$. The precision $\frac{1}{\sigma_{\nu i}^2}$ depends on i , meaning that some families (the most socially advantaged ones) get more information on school quality from the RAR signal than others. A RAR school gets additional resources Δ_j that may increase school quality: $\tilde{q}_j = q_j + \Delta_j$. Families do not observe Δ_j , and their beliefs are supposed to be normally distributed: $C_i(\Delta_j|X, q_j, S_j = 1) \sim \mathcal{N}(\Delta_{ij}, \sigma_{\Delta i}^2)$. Each family updates schools expected qualities, which become

$$\mathbb{E}_i(\tilde{q}_j|X_j, S_j = 1) = X_j'\beta_i + \Delta_{ij} - \sigma_{qi} \frac{\phi}{\Phi} \left(\frac{X_j'\beta_i}{\sqrt{\sigma_{qi}^2 + \sigma_{\nu i}^2}} \right)$$

for a RAR school, and

$$\mathbb{E}_i(\tilde{q}_j | X_j, S_j = 0) = X_j' \beta_i + \sigma_{qi} \frac{\phi}{(1 - \Phi)} \left(\frac{X_j' \beta_i}{\sqrt{\sigma_{qi}^2 + \sigma_{\nu i}^2}} \right)$$

for a non RAR school.

Suppose that family i was indifferent between school A and school B before receiving the RAR signal, so that $X_A' \beta_i = X_B' \beta_i$. Let us now assume that school A is RAR. Family i now chooses school A over school B if and only if:

$$\Delta_{iA} > \sigma_{qi} \left[\frac{\phi}{\Phi} \left(\frac{X_A' \beta_i}{\sqrt{\sigma_{qi}^2 + \sigma_{\nu i}^2}} \right) + \frac{\phi}{(1 - \Phi)} \left(\frac{X_B' \beta_i}{\sqrt{\sigma_{qi}^2 + \sigma_{\nu i}^2}} \right) \right]$$

School A is chosen if Δ_{iA} is sufficiently large, that is, if family i believes that additional resources from the RAR program will significantly increase school quality. On the contrary, if Δ_{iA} is close to zero, or even negative, then family i will choose school B . A family observing school quality perfectly *ex ante* ($\sigma_{qi} = 0$) would lend no weight to the RAR status, other than through the supposed increase in quality Δ_{iA} . If family i believes that the RAR status is not very informative, i.e. $\sigma_{\nu i}$ is large, then it will only care about Δ_{iA} . On the contrary, families with a small $\sigma_{\nu i}$ would lend more weight to the RAR status, and would more likely choose school B instead of A . Furthermore, if family i anticipates that socially more advantaged families are likely not to choose RAR schools, then it may expect a negative Δ_{iA} (because of less good pairs for instance), and it is more likely to choose school B instead of A .

1.3 The data

To analyze the effect of the program on pupils' school choice and academic achievement, we use exhaustive micro-level data provided by the statistical service of the French Ministry of Education, both at the pupil and school levels.

First, we use annual exhaustive individual data sets of French secondary education pupils (called "*fichiers anonymisés d'élèves pour la recherche et les études*" or FAERE). We focus on pupils entering junior high school (6th grade) in 2006 and 2007, that is, the first two cohorts of pupils affected by the RAR program. Pupils entering 6th grade later are not taken into account for two reasons; first, at the time of this study, their scores at the "*Brevet*" national exam and their situations four to five

years after entering 6th grade were unknown; second, the program may have had long term effects on location and primary school choices, which would bias our analysis. These data come from administrative sources and gather some information on pupils: we observe their sexes, ages, origins, their family backgrounds through their parents' occupations, and whether or not they benefit from a scholarship. These data were matched with the exhaustive "*Brevet*" national exam data set.⁹ The situation of pupils just after the end of junior high school is also observed. We know which junior high schools they attend in 6th grade, whether these are public or private schools, and whether these are RAR schools or not. We observe which primary schools they attended the preceding year, with their exact geographic locations.

A second source of data comes from an exhaustive data set at the school level in which we observe every mainland France public junior high school with their exact geographic location.¹⁰ For each of these schools, we observe the proportion of pupils from disadvantaged backgrounds⁵ and the proportion of repeaters when entering 6th grade as evaluated during the 2004-2005 school year, i.e. we observe whether or not each junior high school was eligible to the RAR program.

Combining those two data sets, we are able to define each pupil's closest public junior high school, using the smallest (Euclidean) distance to his or her primary school. Because we don't know the exact location of pupils' homes, we approximate their location by the location of their primary school. Note that we also don't know their catchment area junior high school; instead, we consider the closest public junior high school to their primary school. We believe that this is a convincing proxy for families' default school options. Moreover, as long as there is no systematic bias in our measure of closest schools, then the only consequence is a loss of precision.

In the following, we will consider two treatment variables. The first one is whether pupils live in the vicinity of a RAR junior high school. The second one is whether pupils are enrolled in a RAR junior high school in 6th grade. Because we observe pupils every schooling year, we could have considered being enrolled in a RAR school the whole time of lower secondary education as an alternative variable. However, the vast majority of pupils (more than 80%) attend the same school during

⁹For every cohort, we observe two consecutive "*Brevet*" sessions, so that the results of pupils who repeated one grade during junior high school are observed. When a pupil was present in both sessions, we kept only the first one.

¹⁰In the school-level data-base, we only consider public junior high schools. However, we do observe pupils enrolled in a private school in the individual-level data-base so that the private sector is not excluded from the analysis.

all lower secondary education. Moreover, the results are similar when we consider this alternative variable.

We restrict our sample to pupils living in mainland France,¹¹ and we thus observe 1 098 636 individuals, with 531 729 entering 6th grade in 2006 and 566 907 in 2007. Among them, 45 376 are living in the vicinity of a RAR junior high school, and 28 517 are enrolled in a RAR junior high school; that is, 3% of the sample. More than 50% of pupils are going to their closest public junior high school, while 27% are attending another public school. Around 20% of 6th grade pupils are attending a private school. Among pupils living near a RAR junior high school, 41% are attending the closest (RAR) public junior high school, 40% are going to another public school, and 19% are attending a private school.

A first descriptive statistics analysis shows that pupils attending a RAR junior high school in 6th grade have poorer academic outcomes than other pupils. 69% of them pass the "*Brevet*" national exam four to five years after entering 6th grade, compared to 87% of non RAR pupils. On average, RAR pupils get a total exam score of 8 over 20, compared to 11 over 20 for non RAR pupils. At the end of junior high school, 43% of pupils who entered a RAR school in 6th grade continue in a general upper secondary education track ("*second cycle général ou technologique*"), compared to 62% of non RAR pupils. Because the RAR program aimed at targeting socially and academically disadvantaged pupils, such differences are not surprising and may come from the fact that RAR pupils are a population that is not directly comparable to non RAR pupils (Table 1.1). For instance, pupils attending a RAR junior high school in 6th grade are older, on average, than non RAR pupils (0.21 year older, that is, 2 months and a half older). On average, they come more often from a disadvantaged background (i.e. with a parent blue-collar, retired blue-collar, retired white-collar, or unemployed) (39 percentage points more), and they benefit more often from a scholarship than other pupils (47 percentage points more). Moreover, living near a RAR junior high school is certainly not exogenous either. Pupils living near a RAR junior high school are very different from other pupils according to their observable characteristics (Table 1.2); on average, they are a bit older than other pupils when entering 6th grade (11.22 years old compared to 11.10 for other pupils, or about 1.5 month older), they are less often born French (95% of them compared to 98% of other pupils), they come much more often from a disadvantaged family

¹¹Observations in overseas "*départements*" are excluded from the analysis, because schools are poorly geocoded.

(56% of them compared to 33%), and they benefit more often from a scholarship (48% of them compared to 20% of other pupils). Furthermore, pupils living near a RAR junior high school live in municipalities where the median households revenue is about 4,500 euros smaller on average than in other pupils' municipalities, where the unemployment rate is about 1 point higher and where the population density is about 2,000 inhabitants per square kilometer higher than in other pupils' municipalities (Table 1.3).

Of course, such differences are not surprising, since the RAR program was aimed at targeting pupils in socially disadvantaged schools. But these differences highlight the fact that a naive comparison between pupils affected by the RAR program and other pupils would lead to a selection bias, since they would be very different populations even in the absence of the program.

Table 1.1 – Individual characteristics in the sample

Individual characteristics of	pupils not	pupils	Test (p-value)
	enrolled in a RAR	enrolled in a RAR	
Male	0.50 0.001	0.49 0.003	0.002
Age at 6th grade	11.10 0.001	11.31 0.005	<0.001
Born French	0.98 0.000	0.93 0.005	<0.001
Disadv. background	0.33 0.002	0.72 0.007	<0.001
Scholarship	0.20 0.001	0.67 0.008	<0.001
Nbr obs	1,070,119	28,517	

Note: Standard errors are clustered at the closest junior high school level.

Lecture: 67% of pupils enrolled in a RAR public junior high school benefit from a scholarship compared to 20% of other pupils. The difference of 47 percentage points is significant at the 1% significance level.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.2 – Individual characteristics in the sample

Individual characteristics of	Individual characteristics of		Test (p-value)
	pupils not living near a RAR	pupils living near a RAR	
Male	0.50 0.001	0.49 0.003	0.002
Age at 6th grade	11.10 0.001	11.22 0.005	<0.001
Born French	0.98 0.000	0.95 0.003	<0.001
Disadv. background	0.33 0.002	0.56 0.008	<0.001
Scholarship	0.20 0.001	0.48 0.009	<0.001
Nbr obs	1,053,260	45,376	

Note: Standard errors are clustered at the closest junior high school level.

Lecture: 48% of pupils living in the vicinity of a RAR public junior high school benefit from a scholarship compared to 20% of other pupils. The difference of 28 percentage points is significant at the 1% significance level.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.3 – Characteristics of pupils' municipality of residence in the sample

Characteristics of municipality of	pupils not	pupils	Test (p-value)
	living near a RAR	living near a RAR	
Median households revenue	28,536 132.8	24,024 376.4	<0.001
Unemployment rate	7.63 0.031	8.67 0.196	<0.001
Population density	1,546 81.7	3,509 307.7	<0.001
Nbr obs	967,563	39,238	

Note: Standard errors are clustered at the municipality (of residence) level.

Lecture: Pupils living in the vicinity of a RAR public junior high school live in municipalities where the median households revenue is 28,536 euros on average compared to 24,024 for other pupils. The difference of 4,512 euros is significant at the 1% significance level.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

1.4 Effect on individual sorting

1.4.1 Empirical strategy

Identification

To assess the role of the RAR program on possible sorting across schools, we analyze the effect on school choice of living near a RAR junior high school. Consider the Rubin causal model framework (Rubin, 1974), where the "treatment" dummy variable T_i^{NEAR} equals 1 if pupil i lives near a RAR junior high school and 0 otherwise. Let us consider Y_i the dummy outcome variable of school choice. Y_i will represent different outcomes corresponding to the different possible school choices: being enrolled in the closest junior high school, being enrolled in a private school, or being enrolled in another public school. For those different variables, $Y_i(0)$ and $Y_i(1)$ are potential outcomes. $Y_i(1)$ is the value of the outcome when individual i lives near a RAR and $Y_i(0)$ is the value of the outcome otherwise. The observed outcome is given by:

$$Y_i = Y_i(0) + T_i^{NEAR}(Y_i(1) - Y_i(0))$$

or:

$$Y_i = \alpha_i + T_i^{NEAR}\beta_i$$

where β_i represents the individual treatment effect: the individual effect on school choice of living near a RAR junior high school.

Because, as explained in the preceding part, living near a RAR school is not exogenous, the estimation of β_i using a classical OLS regression is biased. It would partly capture the fact that the population of pupils living near a RAR junior high school would have different potential outcomes than other pupils, even in the absence of treatment. Our identification strategy consists of using the eligibility thresholds to assess causality. More precisely, the principle of our identification strategy is the following: pupils living near a public junior high school that is above the eligibility thresholds have a higher probability that their closest (default option) school is a RAR. Some pupils would thus be treated exogenously, because their closest public junior high school is above the threshold. We thus use a regression discontinuity framework.

In a preceding work, Beffy and Davezies (2013) showed that the probability for a school to enter the RAR program increases discontinuously above the 10% and 67% eligibility thresholds. Not every school above the thresholds entered the program,

while some schools below the thresholds did enter it, so the discontinuity is fuzzy. We use this result to construct a fuzzy regression discontinuity analysis at the pupil level.

Figure 1.2 highlights two clear discontinuities in the individual probability that the closest public junior high school is a RAR around the 10% threshold of repeaters in the nearest public junior high school, and around the 67% of pupils from disadvantaged backgrounds in the nearest public junior high school. Under the assumption that pupils living near a junior high school just below and just above the eligibility thresholds are similar, then any discontinuity in the individual outcome around the thresholds may be interpreted as a causal effect of the proximity of a RAR school.

We consider the two discontinuities separately, and thus, we analyze the effect of the treatment on two sub-populations around the thresholds. Figure 1.3 illustrates the two regions we will consider for a given bandwidth around the thresholds. In other words, we will estimate two effects: the effect of the treatment on school choice around the 10% threshold of repeaters in the nearest school (discontinuity d^L in Figure 1.3) and the effect of the treatment on school choice around the 67% threshold of pupils from disadvantaged backgrounds in the nearest school (discontinuity d^F).¹²

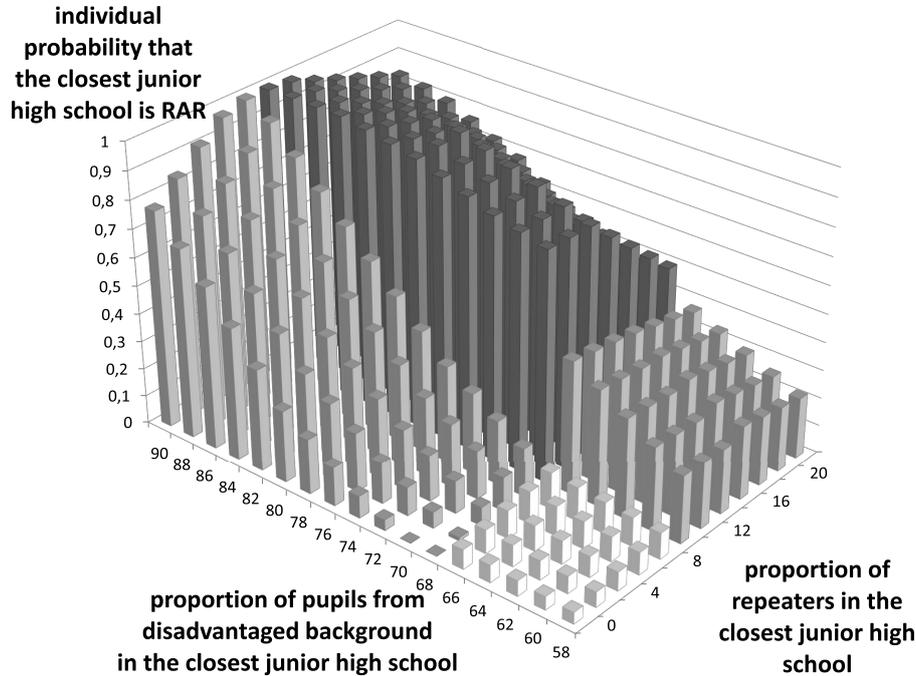
Let us define for each individual i Z_i^L the proportion of pupils who have repeated twice or more in pupil i 's nearest public junior high school, and Z_i^F the proportion of pupils from disadvantaged background in pupil i 's nearest public junior high school. The individual probability to live near a RAR is discontinuous in (Z^L, Z^F) at the thresholds $c^L = 10\%$ and $c^F = 67\%$, respectively. Assignment to treatment is not deterministic; not all units move from non-participation to participation above the threshold, but the probability jumps discontinuously at the threshold (fuzzy design).

13

¹²We could alternatively consider both discontinuities at the same time. In this case, we would consider a two-dimensional assignment "boundary" and compare pupils just above and just below this assignment boundary. However, our preferred specification is to make two separate estimations, because comparing pupils below and above the boundary might lead us to compare pupils from very different neighborhood types (for instance, one in the North-West quarter of the (Z^L, Z^F) graph, and another in the South-East quarter).

¹³At this stage, it is important to keep in mind that, because we don't observe pupils' residence exact location, we may wrongly assign some pupils to a junior high school that is not their closest school. However, for each school, the proportion of repeaters and the proportion of socially disadvantaged pupils are perfectly observed. If some pupils are mismatched with a closest school, then we are in a case of contaminated data on the running variable (Horowitz and Manski, 1995).

Figure 1.2 – Individual probability that the nearest junior high school is RAR



Note: The graph presents on the x-axis, the proportion of pupils from disadvantaged backgrounds in the nearest junior high school; on the z-axis, the proportion of pupils who have repeated twice or more in the nearest junior high schools; and on the y-axis, the individual probability that the nearest junior high school is RAR (estimated by a local linear regression in each quadrants).

Source: MEN-MESR DEPP, FAERE 2006 and 2007

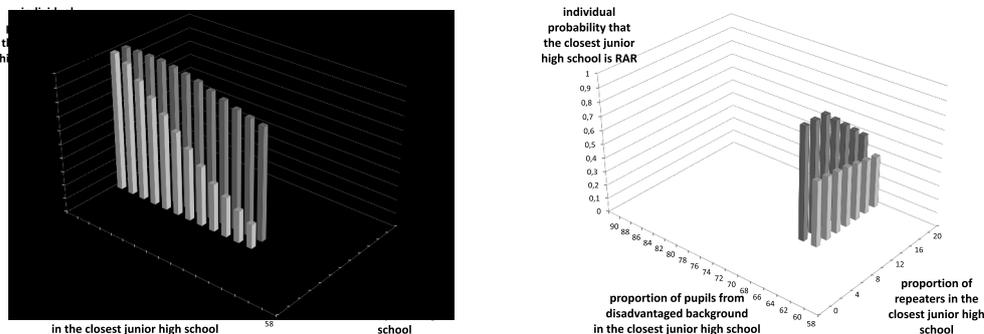
These discontinuities identify the existence of "*compliers*" (Hahn, Todd, and Klaauw, 2001; Imbens and Angrist, 1994; Imbens and Lemieux, 2008), i.e. pupils who live near a RAR junior high school (respectively not) just because their closest school is above (respectively just below) the threshold. Those pupils are treated (or not) exogenously.¹⁴

For the population of compliers, the local average treatment effect at the cutoff

As long as the occurrence of misclassifications is exogenous, then the estimates are consistent. Note that we are not in a situation in which a continuous measurement error affects every units and challenges identification (Davezies and Le Barbanchon, 2014; Pei, 2011; Yu, 2012).

¹⁴More formally, compliers are individuals such that $\lim_{z \uparrow c} T_i^{NEAR}(x) = 0$ and $\lim_{z \downarrow c} T_i^{NEAR}(x) = 1$. This corresponds to the subgroup of individuals for whom treatment changes discontinuously at the threshold.

Figure 1.3 – Illustration of a window around each discontinuity



Source: MEN-MESR DEPP, FAERE 2006 and 2007

values is:

$$LATE_j = E(Y(1) - Y(0)|Z^j = c^j), j \in \{L, F\}$$

Note that this identifies the average treatment effect only locally, for pupils who are treated because their nearest school is just above the threshold and who would not have been treated otherwise.

Under the assumption that $(z^F, z^L) \mapsto \mathbb{E}(Y(j)|Z^F = z^F, Z^L = z^L)$ is continuous for $j = 0, 1$, these parameters are identified by the Wald ratio of the jump in the conditional expectation of the outcome to the jump in the conditional expectation in the treatment indicator:

$$LATE_J = \frac{\lim_{z^J \downarrow c^J} \mathbb{E}(Y|Z^J = z^J, Z^K \in I_K) - \lim_{z^J \uparrow c^J} \mathbb{E}(Y|Z^J = z^J, Z^K \in I_K)}{\lim_{z^J \downarrow c^J} \mathbb{E}(T|Z^J = z^J, Z^K \in I_K) - \lim_{z^J \uparrow c^J} \mathbb{E}(T|Z^J = z^J, Z^K \in I_K)},$$

$(J, K) \in \{(L, F), (F, L)\}$

The assumption that $(z^F, z^L) \mapsto \mathbb{E}(Y(j)|Z^F = z^F, Z^L = z^L)$ is continuous is not directly testable. However, it means that pupils living near schools just above and just below the thresholds are similar. Thus, we can at least compare mean values between these two sub-populations with respect to observable characteristics in the data. For variables which are not correlated to Z , we expect mean values to be close if pupils living near schools above and below the threshold are similar.

Tables 1.4 and 1.5 compare the sub-population of pupils whose closest school is just above the threshold with the sub-population of pupils whose closest school is just below the threshold, with respect to every individual characteristic observed in the data. The joint significance test is presented, and tests whether the two sub-

populations are significantly different (according to observed characteristics). As a reminder, the first column of both tables compares pupils whose nearest junior high school is a RAR and pupils whose nearest junior high school is not a RAR. Again, this column shows that these two populations differ significantly with respect to all observable characteristics. In the three following columns, the same mean differences are computed for the sub-sample of observations lying just around the threshold (for different bandwidths) in order for the results to be comparable with the three last columns. Note that significant differences still hold when we focus on observations just around the thresholds. Then the three remaining columns compare pupils living near a public junior high school that is just below or just above the eligibility thresholds, for different windows around the thresholds. These differences are almost always jointly not significant (in 5 cases out of 6). Tables 1.20 and 1.21 in Appendix 1.C provide the same evidence on pupils' municipality of residence.

Our identification strategy relies on the similarity between pupils living near schools just below and just above the eligibility thresholds. According to these descriptive statistics, they do not differ much with respect to observable characteristics. Although the identifying assumption cannot be formally tested, this provides empirical support to the validity of our approach.

Estimation

We estimate the local average treatment effects with a two-stage least square estimator around the thresholds (Hahn, Todd, and Klaauw, 2001). More precisely, we estimate the following equation by two-stage least squares:

$$Y_i = \alpha + \beta T_i^{NEAR} + \gamma' V_i^j + \varepsilon_i \quad (1.1)$$

where T_i^{NEAR} is considered endogenous and instrumented by $\mathbf{1}\{Z_i^j > c^j\}$, $j \in \{L, F\}$.

We include the additional covariates : $V_i^j = \begin{pmatrix} \mathbf{1}\{Z_i^j < c^j\}(Z_i^j - c^j) \\ \mathbf{1}\{Z_i^j \geq c^j\}(Z_i^j - c^j) \end{pmatrix}$, $j \in \{L, F\}$ in order to allow the slope coefficient to be different on each side of the cutoff. This aims at limiting the asymptotic bias of non parametric estimators (Imbens and Lemieux, 2008).

As mentioned earlier, we estimate the model separately around each discontinuity. More precisely, we estimate equation (1.1) by 2SLS:

1. for observations such that $Z^L \in [10 - h_L, 10 + h_L]$ and¹⁵ $63 \leq Z^F \leq 90$

¹⁵We restrict the sample around Z^F because there is no discontinuity in Z^L outside of this interval

Table 1.4 – Individual characteristics around discontinuity d^L

	Mean comparison of:						
	Pupils living near a RAR vs not				Pupils above vs below d^L		
	Total	h=0.5	h=1	h=2	h=0.5	h=1	h=2
Male	-0.00 (0.00)	-0.01 (0.02)	0.00 (0.01)	0.00 (0.01)	-0.01 (0.02)	0.01 (0.01)	-0.00 (0.01)
Age at 6th grade	0.01*** (0.00)	0.01 (0.02)	0.03* (0.02)	0.03** (0.01)	0.01 (0.02)	0.01 (0.02)	-0.00 (0.01)
Born French	-0.06*** (0.01)	-0.06 (0.07)	-0.15** (0.06)	-0.06 (0.05)	-0.06 (0.07)	-0.10 (0.10)	-0.09 (0.06)
Disadvantaged background	0.02*** (0.00)	-0.05 (0.03)	-0.04 (0.03)	-0.01 (0.02)	-0.05 (0.03)	-0.04 (0.03)	-0.04** (0.02)
Scholarship	0.06*** (0.00)	0.13** (0.05)	0.13*** (0.04)	0.14*** (0.03)	0.13** (0.05)	-0.01 (0.06)	-0.01 (0.03)
Test (pvalue)	0.000	0.183	0.000	0.000	0.183	0.709	0.182
Nbr obs	1,071,395	1,713	6,473	15,303	1,713	6,473	15,303
Nbr clusters	9,931	24	68	154	24	68	154

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the closest junior high school level. Mean differences are estimated for different bandwidth of size h_L around the 10% threshold.

Lecture: Pupils who benefit from a scholarship have a higher probability to live in the vicinity of a RAR junior high school (6 percentage points more) than non scholarship pupils, all other characteristics being equal. This difference is significant at the 1% level and remains significant for different sub-samples of observations lying just around the 10% threshold.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

2. for observations such that $Z^F \in [67 - h_F, 67 + h_F]$ and $Z^L \geq 10$

where h_L and h_F are bandwidths around the thresholds.¹⁶

In a second part, to assess possible heterogeneous effects in the population, we will consider that the treatment effect can be different across some sub-populations. To do so, we interact the treatment variable with some covariates. Let us note X_i

(Figure 1.2). This interval was chosen so as to have a genuine discontinuity in Z^L without losing too many observations.

¹⁶We tested for the robustness of our estimates to the bandwidth choice. For each discontinuity, three fixed bandwidths are used. We also estimated the effects using the optimal bandwidth (Imbens and Kalyanaraman, 2012) for each outcome, twice and half the size of the optimal bandwidth (results are available on request). The results are robust to the bandwidth choice.

Table 1.5 – Individual characteristics around discontinuity d^F

	Mean comparison of:						
	Pupils living near a RAR vs not				Pupils above vs below dF		
	Total	h=2	h=4	h=8	h=2	h=4	h=8
Male	-0.00 (0.00)	-0.00 (0.02)	0.00 (0.01)	0.00 (0.01)	-0.02 (0.02)	0.01 (0.01)	0.00 (0.01)
Age at 6th grade	0.01*** (0.00)	-0.03 (0.04)	-0.00 (0.02)	0.00 (0.01)	-0.05 (0.03)	-0.01 (0.02)	-0.01 (0.01)
Born French	-0.06*** (0.01)	-0.08 (0.10)	-0.01 (0.10)	-0.08* (0.05)	0.22** (0.09)	0.11 (0.10)	-0.04 (0.05)
Disadvantaged background	0.02*** (0.00)	-0.09* (0.05)	-0.03 (0.04)	0.02 (0.02)	-0.02 (0.04)	-0.01 (0.03)	0.04** (0.02)
Scholarship	0.06*** (0.00)	-0.02 (0.05)	0.13** (0.05)	0.08*** (0.03)	-0.13*** (0.04)	0.04 (0.05)	0.03 (0.03)
Test (pvalue)	0.000	0.529	0.022	0.016	0.052	0.775	0.162
Nbr obs	1,071,395	2,475	5,416	14,897	2,475	5,416	14,897
Nbr clusters	9,931	30	58	150	30	58	150

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the closest junior high school level. Mean differences are estimated for different bandwidth of size h_F around the 67% threshold.

Lecture: Pupils one year older have a 1 percentage point higher probability to live in the vicinity of a RAR junior high school, all other characteristics being equal. This coefficient is significant at the 1% level and turns not significant in the sub-samples of observations lying just around the 67% threshold.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

as the dummy variable that corresponds alternatively to the scholarship status of pupil i , or being from a disadvantaged background or not. The regression equation becomes:

$$Y_i = \alpha + \beta_1 T_i^{NEAR} \times (1 - X_i) + \beta_2 T_i^{NEAR} \times X_i + \beta_3 X_i + \gamma_1' V_i + \gamma_2' (V_i \times X_i) + \varepsilon_i \quad (1.2)$$

where $T_i^{NEAR} \times (1 - X_i)$ and $T_i^{NEAR} \times X_i$ are instrumented by $\mathbb{1}\{Z^j > c^j\} \times (1 - X_i)$ and $\mathbb{1}\{Z^j > c^j\} \times X_i$. In that case, we are interested in the terms β_1 and β_2 , which represent the effect of living near a RAR school on the sub-population of pupils such that $X = 0$, and on the sub-population of pupils such that $X = 1$, respectively. In order to test for heterogenous effects between those two sub-populations, we will test for $H_0 : \beta_1 = \beta_2$.

1.4.2 Results

We analyze the impact of the RAR program on each possible school choice that pupils face when entering 6th grade, i.e. enrolling in the nearest school, in another public school, or in a private school. Table 1.6 presents the results of the two-stage least square estimations around each discontinuity. The effects are estimated for different bandwidths h around each threshold.

The first stage estimates are consistent with the validity of our strategy; whatever the bandwidth, the coefficients corresponding to $\mathbb{1}\{Z_i^L > 10\}$ and $\mathbb{1}\{Z_i^F > 67\}$, respectively, are highly significant in the first stage regression. Living near a public junior high school where the proportion of repeaters is above the 10% threshold significantly increases by 79 to 84 percentage points the individual probability that the closest junior high school is a RAR. This corresponds to the proportion of pupils whose closest junior high school is a RAR exogenously (due to the fact that their closest school is just above the threshold) and who would otherwise not live close to a RAR. The effect of living near a RAR is then estimated on those pupils (i.e. the *compliers*) in the second stage. The same results hold when the proportion of pupils from disadvantaged backgrounds in the nearest school is above the 67% threshold. Moreover, F-statistics are always well above the rule of thumb value of 10.

Note that for a sufficiently small bandwidth ($h = 0.5$), the first stage estimate around the 10% threshold is 1, meaning that assignment to treatment is deterministic above the threshold: all schools above the threshold are RAR, and no school below the threshold is RAR. In that case, the proportion of *compliers* is 100%, and we no longer need a two-stage least square estimation; we can directly regress the outcome variable on being above the threshold. This is displayed in the first column of Table 1.6. The remaining columns present the second stage estimates for different larger sizes of the bandwidth. For sake of clarity, we only present the coefficient corresponding to the treatment dummy T^{NEAR} , i.e. living near a RAR junior high school, in the table.

Effects on school choice

When significant, these estimates are negative for enrollment in the nearest school and positive for enrollment in a private school. If one considers the first column of Table 1.6, the results suggest that living near a RAR junior high school decreases the probability to attend this school by 30 percentage points and increases the

probability to attend a private school by 22 percentage points, for pupils who are treated exogenously because their closest school is just above the 10% threshold of repeaters. The fifth column of Table 1.6 suggests that living near a RAR junior high school increases the probability to enroll in a private school by 38 percentage points, for pupils who are treated exogenously because their closest school is just above the 67% threshold of pupils from a disadvantaged background. These results become nonsignificant when we increase the size of the bandwidth. Since the bias is expected to increase with the size of the bandwidth, however, our preferred specification should be that with the smallest bandwidth.

These results mean that living near a RAR junior high school tends to decrease the individual probability to attend the default option junior high school (the closest one) and to increase the individual probability to attend a private school, for the population of *compliers*. This suggests that individuals tend to avoid schools that enter the RAR program by enrolling in the private sector.

Heterogeneous effects with respect to individual social characteristics

In order to better understand these averaged results, we turn to the estimation of heterogeneous effects in the population of *compliers*. Table 1.7 presents the results when we allow living near a RAR school to have differentiated effects in the sample according to observed social characteristics. More precisely, we consider that the effect of living near a RAR school can be different between pupils who benefit from a scholarship and pupils who don't. Then we consider that the treatment effect can be differentiated according to whether or not pupils come from a disadvantaged background. For sake of simplicity, we present only the results estimated around discontinuity d^L (the results around discontinuity d^F are displayed in Table 1.22 in Appendix 1.D), and only the coefficients corresponding to the treatment dummy interacted with $(1 - X)$ and with X respectively, for the two different covariates just mentioned.

For the sub-population of pupils who don't benefit from a scholarship, we find that living near a RAR junior high school due to the 10% threshold decreases the probability to attend the closest school (by 13 to 99 percentage points, depending on the bandwidth size). This effect is highly significant for the two smallest bandwidths. There is no significant effect for the sub-population of pupils who benefit from a scholarship. The difference of the effect between scholarship and non-scholarship pupils is significant at the 5% level for all bandwidths. The estimated effect is

quantitatively the same for pupils who don't come from a disadvantaged background. Again, there is no significant effect for pupils who come from a disadvantaged background. The difference between those two sub-populations is also significant at the 5% level.

The effect on the probability to enroll in another public school is not that clear, considering that the sign of the coefficient changes when the size of the bandwidth increases.

The probability to enroll in a private school increases significantly when pupils live near a RAR school, due to the 10% threshold, for the sub-population of pupils who don't benefit from a scholarship and for the sub-population of pupils who don't come from a disadvantaged background. In particular, for the sub-population of pupils who don't come from a disadvantaged background, living near a RAR junior high school due to the 10% threshold significantly increases the probability to enroll in the private sector by 27 to 63 percentage points, whereas we find no significant effect for the sub-population of pupils from a disadvantaged background.

These results highlight the mechanisms underlying individual school choice. It seems that the most socially advantaged individuals tend to avoid schools that enter the RAR program more than other pupils, by enrolling instead in the private sector.

Why do socially advantaged families try and avoid treated schools more than other families? First, it may be that they are financially less constrained, so that they can afford the fees to send their child(ren) in a private school, for instance, whereas other families cannot. Second, and more interestingly, they may have better information about the educational system. As an example, they are likely to better know the procedure to ask for another school than the catchment area one. Or they may get more information from the treatment status than other families. In order to analyze these hypotheses, let us assume that parents being teachers themselves have a better information about the educational system than other parents. If parents who are close to the educational system can get more information on schools and on how to change school,¹⁷ then the effect of living in the vicinity of a treated school should be different for children of teachers. Let us construct a dummy variable which equals one if one of the parents is a teacher, and zero otherwise. About 7% of pupils in the data have at least one parent being a teacher. The first part of Table 1.8 presents

¹⁷Until 2007, teachers could ask for their child(ren) to be enrolled in the school where they worked. From the 2007 school year, this was not a motive for asking for another school than the catchment area one anymore. However, there is no difference between the 2006 cohort and the 2007 cohort in the probability for children of teachers to avoid treated schools.

the effect of living in the vicinity of a treated school interacted with the teacher dummy. The results of the equality test suggest that living near a treated school decreases more the probability to enroll in the nearest school for children of teachers than for other children. The treatment effect on enrollment in another school, and in a private school, is also higher for children of teachers than for other children, though the difference is not significant for every bandwidths h . However, comparing teachers to all other professions may not be enough to conclude. They may be better informed, but they are also more educated and have a higher household income on average than the rest of the population. To better distinguish whether this difference is due to an informational advantage, or to a higher economic and cultural capital, the second part of Table 1.8 compares the treatment effect for children of teachers, and for children of executives only. The assumption is that both populations have an educational and economic advantage, but teachers have a better knowledge of the educational system. The results show that the difference in the effect of living near a treated school on school choice between teachers and other executives is much smaller. More precisely, both teachers and executives have a higher probability to send their children in a private school when the closest school is a RAR, but the difference is not significant. However, children of teachers have a higher probability to enroll in another public school when the closest school is a RAR. On the contrary, children of executives do not enroll significantly more in another public school when the closest school is RAR, the coefficient is even negative and marginally significant on the larger bandwidth. This may suggest that part of parents' strategies to avoid treated schools is due to an informational advantage.

Table 1.6 – Estimation of the effect of living near a RAR junior high school on school choice

	RD around the 10% threshold			RD around the 67% threshold		
	h=0.5	h=1	h=2	h=2	h=4	h=8
Second stage						
<i>Y=Enrollment in the nearest school</i>						
<i>T</i>	-0.30*** (0.07)	-0.46*** (0.17)	0.05 (0.13)	0.08 (0.23)	-0.29 (0.20)	-0.19 (0.16)
<i>Y=Enrollment in another public school</i>						
<i>T</i>	0.08 (0.06)	0.12 (0.15)	-0.15 (0.14)	-0.01 (0.19)	-0.08 (0.16)	0.08 (0.15)
<i>Y=Enrollment in a private school</i>						
<i>T</i>	0.22** (0.08)	0.34 (0.21)	0.10 (0.11)	-0.07 (0.13)	0.38* (0.20)	0.11 (0.16)
First stage						
$\mathbb{1}\{Z^j > c^j\}$	1.00*** (0.00)	0.84** (0.32)	0.79*** (0.21)	0.74** (0.31)	0.63*** (0.21)	0.44*** (0.15)
F-stat	.	572	1,624	415	1,000	1,197
Nbr obs	1,778	6,671	15,441	2,548	5,602	15,022
Nbr clusters	24	68	150	30	58	144

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the closest junior high school level. Two-stage least squares are estimated for different bandwidths of size h around each threshold.

Lecture: Pupils living near a RAR junior high school exogenously, due to the fact that their closest public junior high school is above the 10% threshold of repeaters, have a 10 to 34 percentage point higher probability to enroll in a private school than pupils whose closest public junior high school is not a RAR exogenously. This difference is significant at the 5% significance level for the smallest window around the 10% threshold.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.7 – Estimation of heterogeneous effects of living near a RAR on school choice (d^L)

Y=Enrollment in	the nearest school			another public school			a private school		
	RD around the 10% threshold								
	h=0.5	h=1	h=2	h=0.5	h=1	h=2	h=0.5	h=1	h=2
<i>X=Scholarship (ref.=no)</i>									
<i>X=0</i>	-0.61*** (0.09)	-0.99*** (0.30)	-0.13 (0.14)	0.22*** (0.06)	0.39*** (0.13)	-0.07 (0.13)	0.40*** (0.12)	0.60** (0.30)	0.20 (0.14)
<i>X=1</i>	0.12 (0.28)	0.17 (0.29)	0.34* (0.20)	-0.18 (0.28)	-0.22 (0.32)	-0.28 (0.20)	0.06** (0.03)	0.05 (0.05)	-0.06 (0.09)
Test (pvalue)	0.030	0.033	0.032	0.163	0.068	0.224	0.006	0.054	0.054
Nbr obs	1,778	6,671	15,441	1,778	6,671	15,441	1,778	6,671	15,441
Nbr clusters	24	68	150	24	68	150	24	68	150
<i>X=Disadvantaged background (ref.=no)</i>									
<i>X=0</i>	-0.64*** (0.12)	-1.04** (0.40)	-0.15 (0.13)	0.25*** (0.09)	0.41*** (0.15)	-0.12 (0.13)	0.39*** (0.12)	0.63* (0.38)	0.27* (0.16)
<i>X=1</i>	0.00 (0.13)	-0.05 (0.17)	0.26 (0.17)	-0.03 (0.11)	-0.07 (0.21)	-0.16 (0.17)	0.03 (0.06)	0.12 (0.09)	-0.09 (0.08)
Test (pvalue)	0.006	0.050	0.022	0.147	0.083	0.747	0.001	0.094	0.008
Nbr obs	1,713	6,473	14,887	1,713	6,473	14,887	1,713	6,473	14,887
Nbr clusters	24	68	150	24	68	150	24	68	150

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the closest junior high school level. Two-stage least squares are estimated for different bandwidths of size h around the threshold.

Lecture: Scholarship pupils who live near a RAR junior high school exogenously, due to the fact that their closest public junior high school is just above the 10% threshold ($h=0.5$), have a 6 percentage point higher probability to enroll in a private junior high school than scholarship pupils whose nearest junior high school is not a RAR exogenously. Non scholarship pupils who live near a RAR junior high school exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, have a 40 percentage point higher probability to enroll in a private junior high school than non scholarship pupils whose nearest junior high school is not a RAR exogenously. The difference between these two estimates is significantly different from zero.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.8 – Estimation of the effects of living near a RAR on school choice according to knowledge of the educational system (d^L)

Y=Enrollment in	the nearest school			another public school			a private school		
	RD around the 10% threshold								
	h=0.5	h=1	h=2	h=0.5	h=1	h=2	h=0.5	h=1	h=2
<i>X=Teacher (ref.=no)</i>									
<i>X=0</i>	-0.23*** (0.06)	-0.40** (0.16)	0.09 (0.13)	0.05 (0.05)	0.09 (0.16)	-0.17 (0.14)	0.19* (0.09)	0.32 (0.20)	0.08 (0.11)
<i>X=1</i>	-1.12*** (0.16)	-1.71*** (0.56)	-0.43* (0.23)	0.45** (0.20)	0.88*** (0.30)	-0.05 (0.22)	0.67*** (0.23)	0.83** (0.41)	0.48** (0.24)
Test (pvalue)	0.000	0.021	0.014	0.052	0.025	0.505	0.047	0.127	0.029
Nbr obs	1,733	6,535	15,471	1,733	6,535	15,471	1,733	6,535	15,471
Nbr clusters	24	68	154	24	68	154	24	68	154
<i>X=Teacher (ref.=executives)</i>									
<i>X=0</i>	-0.72*** (0.19)	-1.01** (0.44)	-0.30** (0.15)	-0.02 (0.18)	0.26 (0.25)	-0.33* (0.19)	0.74*** (0.14)	0.76** (0.36)	0.63*** (0.23)
<i>X=1</i>	-1.12*** (0.16)	-1.71*** (0.56)	-0.43* (0.23)	0.45** (0.20)	0.88*** (0.30)	-0.05 (0.22)	0.67** (0.24)	0.83** (0.41)	0.48** (0.24)
Test (pvalue)	0.079	0.009	0.421	0.091	0.007	0.075	0.795	0.765	0.217
Nbr obs	211	997	2,194	211	997	2,194	211	997	2,194
Nbr clusters	23	65	146	23	65	146	23	65	146

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the closest junior high school level. Two-stage least squares are estimated for different bandwidths of size h around the threshold.

Lecture: Children of teachers who live near a RAR junior high school exogenously, due to the fact that their closest public junior high school is just above the 10% threshold ($h=0.5$), have a 67 percentage point higher probability to enroll in a private junior high school than children of teachers whose nearest junior high school is not a RAR exogenously. Other pupils who live near a RAR junior high school exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, have a 19 percentage point higher probability to enroll in a private junior high school than non scholarship pupils whose nearest junior high school is not a RAR exogenously. The difference between these two estimates is significantly different from zero.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Heterogeneous effects with respect to local private school supply

Because most of parental strategies to avoid treated schools seem to be driven by pupils enrolling in the private sector, the effect is expected to differ according to the size of the local private supply. As an measure of local private school supply, we defined the number of private junior high schools within 10 kilometers around pupils' preceding primary school. To test whether the effects depend on the alternative supply provided by the private sector in families' neighborhoods, this measure was interacted with the treatment dummies in the model. Table 1.9 gives the result around discontinuity d^L . As before, these results highlight that the probability to avoid the nearest school when it is treated is higher for the sub-population of pupils who do not benefit from a scholarship ($X = 0$) than for pupils who benefit from a scholarship ($X = 1$). Moreover, this effect is even larger when the number of private schools in the neighborhood is higher. Every additional private school in the neighborhood decreases the probability for non scholarship pupils to enroll in the nearest treated school by 0.5 to 1.2 percentage point, and increases the probability to enroll in a private school by 1.1 to 1.2 percentage point. For the sub-population of pupils not coming from a disadvantaged background, every additional private school in the neighborhood decreases the probability to enroll in the nearest school by 0.5 to 1.5 percentage point, and increases the probability to enroll in a private school by 1.2 to 1.3 percentage point.

Table 1.9 – Estimation of the effects of living near a RAR on school choice according to private school supply (d^L)

Y=Enrollment in	the nearest school			another public school			a private school		
	RD around the 10% threshold								
	h=0.5	h=1	h=2	h=0.5	h=1	h=2	h=0.5	h=1	h=2
<i>X=Scholarship (ref.=no)</i>									
<i>X=0</i>	-0.475*** (0.095)	-0.925*** (0.355)	0.121 (0.173)	0.255*** (0.069)	0.537*** (0.132)	-0.106 (0.197)	0.220** (0.103)	0.387 (0.324)	-0.015 (0.161)
<i>X=1</i>	0.863** (0.314)	0.313 (0.308)	0.258 (0.165)	-0.879*** (0.313)	-0.281 (0.332)	-0.220 (0.161)	0.015 (0.025)	-0.032 (0.060)	-0.038 (0.083)
<i>(X=0) × pri</i>	-0.009*** (0.002)	-0.005 (0.004)	-0.012*** (0.003)	-0.002 (0.004)	-0.007** (0.003)	0.001 (0.006)	0.011*** (0.004)	0.012*** (0.004)	0.011** (0.005)
<i>(X=1) × pri</i>	0.086*** (0.021)	0.019 (0.031)	0.008 (0.009)	-0.089*** (0.021)	-0.030 (0.037)	-0.004 (0.007)	0.003** (0.001)	0.011* (0.006)	-0.004 (0.004)
Test (pvalue)	0.000	0.107	0.038	0.001	0.046	0.493	0.003	0.145	0.008
Nbr obs	1,778	6,503	15,700	1,778	6,503	15,700	1,778	6,503	15,700
Nbr clusters	24	68	154	24	68	154	24	68	154
<i>X=Disadvantaged background (ref.=no)</i>									
<i>X=0</i>	-0.479*** (0.122)	-1.028** (0.493)	0.175 (0.174)	0.284*** (0.095)	0.612*** (0.177)	-0.188 (0.212)	0.195* (0.103)	0.416 (0.428)	0.013 (0.180)
<i>X=1</i>	-0.016 (0.239)	0.007 (0.192)	0.240 (0.159)	-0.212 (0.166)	-0.109 (0.257)	-0.138 (0.145)	0.228* (0.116)	0.101 (0.125)	-0.102 (0.087)
<i>(X=0) × pri</i>	-0.011*** (0.002)	-0.005 (0.005)	-0.015*** (0.003)	-0.002 (0.002)	-0.008*** (0.003)	0.003 (0.006)	0.013*** (0.003)	0.013*** (0.004)	0.012*** (0.005)
<i>(X=1) × pri</i>	-0.024 (0.044)	0.019 (0.028)	0.005 (0.007)	-0.024 (0.018)	-0.034 (0.044)	-0.004 (0.006)	0.048 (0.028)	0.015 (0.017)	-0.001 (0.004)
Test (pvalue)	0.105	0.217	0.015	0.121	0.058	0.698	0.158	0.541	0.000
Nbr obs	1,713	6,313	15,143	1,713	6,313	15,143	1,713	6,313	15,143
Nbr clusters	24	68	154	24	68	154	24	68	154

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the closest junior high school level. Two-stage least squares are estimated for different bandwidths of size h around the threshold.

Lecture: Scholarship pupils who live near a RAR junior high school exogenously, due to the fact that their closest public junior high school is just above the 10% threshold ($h=0.5$), have a 1.1 percentage point higher probability to enroll in a private junior high school than scholarship pupils whose nearest junior high school is not a RAR exogenously, and every additional private school in the neighborhood increases this effect by 0.3 percentage point. Non scholarship pupils who live near a RAR junior high school exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, have a 22 percentage point higher probability to enroll in a private junior high school than non scholarship pupils whose nearest junior high school is not a RAR exogenously, and every additional private school in the neighborhood increases this effect by 1.1 percentage point. The difference between both pairs of estimates is jointly significantly different from zero.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

1.5 Effect on individual academic achievement

1.5.1 Empirical strategy

Identification

Let us now turn to the analysis of pupils academic achievement. Following the same idea as in the preceding part, we use the eligibility criteria in pupil's closest school to identify the causal effect of the RAR program on individual performance at the "Brevet" national exam and schooling situation after junior high school.

Let us define T_i^{RAR} the new treatment variable, which equals 1 if pupil i is enrolled in a RAR junior high school in 6th grade and 0 otherwise. The variable Y_i is a measure of pupil i educational achievement (passing the "Brevet" national exam, pupil's score in math, and in French, etc.). $Y_i(1)$ is the potential outcome if pupil i is enrolled in a RAR school in 6th grade and $Y_i(0)$ is the counter-factual potential outcome otherwise. The observed educational outcome is:

$$Y_i = Y_i(0) + T_i^{RAR}(Y_i(1) - Y_i(0))$$

or:

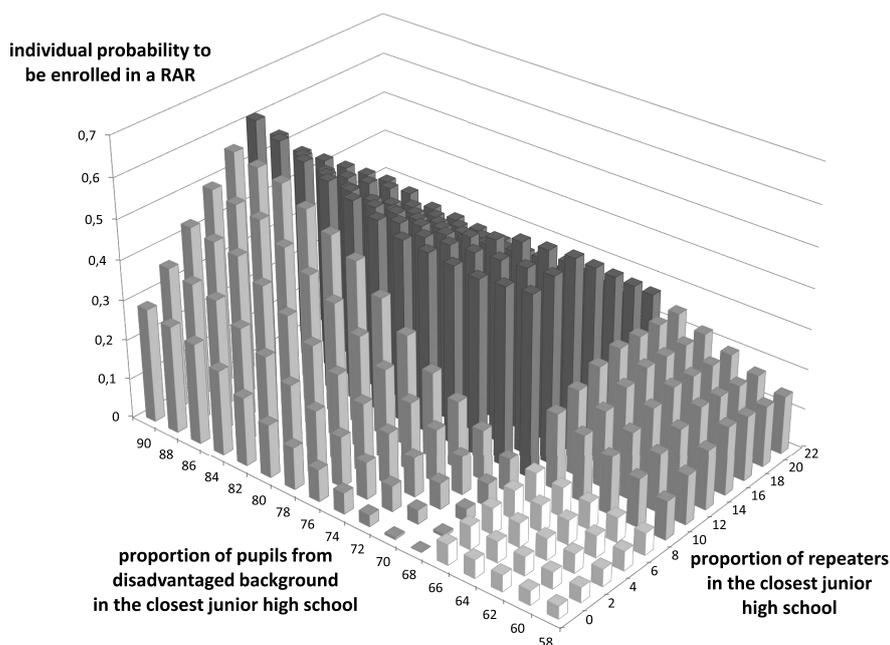
$$Y_i = \alpha_i + T_i^{RAR}\beta_i$$

A naive estimation of the parameter of interest β_i would be distorted by two types of bias. The first one is due to the fact that, as shown in the descriptive statistics, pupils enrolled in a RAR are not directly comparable to other pupils, because RAR schools were selected on the basis of a social and academic disadvantage. An OLS estimate would thus partly capture the fact that RAR pupils would perform worse than other pupils, even without participating in the program. A second source of bias is due to the fact that, as shown in section 1.4.2, pupils selected themselves in (or out of) the treatment, through school choice. Because these school choices are not random (second part of section 1.4.2), the population of pupils enrolled in a RAR is made even less comparable to pupils who are not enrolled in a RAR. In other words, the estimation of β_i by OLS would thus be the result of (at least) three effects; the "true" effect of the program on pupils' performance; the selection effect of ex-ante differences between pupils who would enroll in a RAR and other pupils; and the additional effect of heterogeneous sorting.

To assess the causal impact of the RAR program on pupils' educational outcomes and estimate only the first effect mentioned, we again make use of the eligibility

scheme, using pupils' closest junior high schools. Figure 1.4 shows the discontinuities in the individual probability to be enrolled in a RAR junior high school in 6th grade.

Figure 1.4 – Individual probability to be enrolled in a RAR junior high school



Note: The graph presents on the x-axis, the proportion of pupils from disadvantaged backgrounds in the nearest junior high school; on the z-axis, the proportion of pupils who have repeated twice or more in the nearest junior high schools; and on the y-axis, the individual probability to be enrolled in a RAR junior high school in 6th grade (estimated by a local linear regression in each quadrants).

Source: MEN-MESR DEPP, FAERE 2006 and 2007

In a preceding work, Beffy and Davezies (2013) found that class size decreased in RAR schools, although this decrease was smaller than expected. They also showed a modification of the population in RAR schools; the proportion of children of blue collars increased, and the proportion of children of white collars decreased. They found that the mean results to the "Brevet" national exam decreased in RAR junior high schools. Since their study is made at the junior high school level, the authors cannot disentangle the "true" effect of the program from sorting effects. Thus an analysis at the school level is not sufficient to analyze the impact of the RAR program.

In this chapter, we propose to make the analysis at the pupils' level in order to

account for the individual sorting effects that we highlighted in section 1.4.2. The principle is similar to the preceding part: the individual probability to be enrolled in a RAR school in 6th grade increases discontinuously when the closest school is above the eligibility thresholds. Some pupils are treated exogenously, only because their closest junior high school is above the thresholds. Under the previously stated assumption that pupils living near schools just above and just below the thresholds are similar, then any discontinuity in the outcomes may be interpreted as a causal impact of the RAR program.

Estimation

As before, the local average treatment effects are estimated by two-stage least squares around the thresholds. For each discontinuity separately, the estimated equation is the following:

$$Y_i = \alpha + \beta T_i^{RAR} + \gamma' V_i^j + \varepsilon_i \quad (1.3)$$

where T_i^{RAR} is instrumented by $\mathbb{1}\{Z_i^j > c^j\}$, and $V_i^j = \begin{pmatrix} \mathbb{1}\{Z_i^j < c^j\}(Z_i^j - c^j) \\ \mathbb{1}\{Z_i^j \geq c^j\}(Z_i^j - c^j) \end{pmatrix}$, for $j \in \{L, F\}$.

Equation 1.3 is estimated separately:

1. for observations such that $Z^L \in [10 - h_{10}, 10 + h_{10}]$ and¹⁸ $67 \leq Z^F \leq 80$,
2. and for observations such that $Z^F \in [67 - h_{67}, 67 + h_{67}]$ and $Z^L \geq 10$

with h_{10} (respectively h_{67}) a bandwidth around the 10% (respectively 67%) threshold.

The estimated effect is a (local) average effect. In a second part, we consider that the effect of the program may be different for different sub-populations in the sample. As before, we consider a second model in which the treatment variable is interacted with some individual characteristics. Let X_i be the dummy variable which equals 1 if pupil i entered 6th grade in 2007 and 0 otherwise (we will alternatively consider a variable for pupil's sex, social origin and scholarship status). The second stage equation becomes:

$$Y_i = \alpha + \beta_1 T_i^{RAR} \times (1 - X_i) + \beta_2 T_i^{RAR} \times X_i + \beta_3 X_i + \gamma_1 V_i^j + \gamma_2 (V_i^j \times X_i) + \varepsilon_i$$

¹⁸Again the sample around Z^F was restricted in order to have both a large number of observations and a large discontinuity in Z^L .

where $T_i^{RAR} \times (1 - X_i)$ and $T_i^{RAR} \times X_i$ are instrumented by $\mathbb{1}\{Z_i^j \geq s^j\} \times (1 - X_i)$ and $\mathbb{1}\{Z_i^j \geq s^j\} \times X_i$ in a first stage. The parameter of interest are β_1 and β_2 , which represent the effect of being enrolled in a RAR for the *compliers* such that $X = 0$, and for the *compliers* such that $X = 1$ respectively. Moreover, $H_0 : \beta_1 = \beta_2$ tests whether the effect of being enrolled in a RAR is different for these two sub-populations.

1.5.2 Results

To analyze the effect on individual educational achievement of the RAR program, we use pupils' academic outcomes at the end of junior high school. In a first part, we estimate the effect on the individual probability to pass the "*Brevet*" national exam of enrollment in a RAR school. This analysis is completed with estimations on grades: final exam grade, exam grade in French, exam grade in mathematics, mean continuous assessment grade, and the average grade obtained by combining these different scores. In a second part, we will consider the effect on pupils' situations five years after entering 6th grade. Three cases will be distinguished: enrollment in a general upper secondary education track ("*seconde générale et technologique*"), enrollment in a vocational upper secondary education track, or enrollment in lower secondary education (meaning that the pupil repeated a grade in junior high school).

First stage estimates are systematically reported in the Tables and confirm the validity of the estimation strategy. For each bandwidth size h_{10} around the threshold, being enrolled in 5th grade in a primary school that is close to a junior high school above the 10% threshold of repeaters significantly increases the individual probability to be enrolled in a RAR junior high school in 6th grade by 35 to 70 percentage points, depending on the bandwidth size. This result holds also for every bandwidth size h_{67} around the 67% threshold of pupils from disadvantaged backgrounds: the individual probability to be enrolled in a RAR in 6th grade significantly increases by 40 to 65 percentage points, depending on the size of the bandwidth. These figures represent the proportion of pupils exogenously affected by the RAR treatment in the sample. The effect of the treatment is estimated for those pupils in the second stage. In the first stage, the Fisher statistics are always above the rule of thumb value of 10, which usually justifies the use of instrumental variable instead of ordinary least squares.

For sake of comparison, the results obtained by OLS are also given, on the same samples as the regressions on discontinuity. This allows to better quantify the bias due to selection and sorting.

Effect on individual academic achievement at the end of junior high school

Pupils enrolled in a RAR school pass less often the "*Brevet*" national exam than other pupils. This difference still holds when the sample is restricted to pupils who attended primary schools that are close to a junior high school just below or just above the thresholds. The probability to pass the exam is 10 to 16 percentage points smaller for RAR pupils (OLS estimation in Tables 1.10 and 1.11). But this difference may be due to the selection and sorting bias: pupils enrolled in RAR schools (even those being near the thresholds) are not directly comparable to non RAR pupils. This is why we need to estimate these differences in academic performances only for the pupils who are enrolled in a RAR school exogenously (RDD estimates in Tables 1.10 and 1.11).

These RDD estimates show that the average treatment effects on this population are a lot smaller than the OLS naive estimates. When the coefficients' sign is negative, the absolute value of the estimate is much closer to zero than the OLS estimate on the same sample. In some cases, the coefficient even changes sign and becomes positive. Most importantly, in all cases, the estimates are not significantly different from zero, meaning that no significant effect of the treatment is detected on the sub-population of pupils treated exogenously.

The results are similar when we consider the probability to pass the exam with honors. RAR pupils get the "*Brevet*" with honors significantly less often than other pupils, even when the sample is restricted to pupils whose nearest junior school is close to the thresholds. But this difference diminishes and turns not significant when only exogenously treated pupils are considered.

These results on the probability to pass the "*Brevet*" exam can be completed with an analysis of pupils' exam scores. The exam consists of two different types of tests. On one hand, the grades obtained during the school year count for 60% of the final grade; the remaining 40% come from a written exam covering three subjects: French, Mathematics, and History and Geography (including Civics). Continuous assessment grades depend on the school evaluation practice. Thus they are likely to be influenced by the fact that the school is RAR or not, especially if notation practices are not independent from the class average performance (Murat, 1998, showed that the smaller the level of pupils, the more they are over-graded). In order to analyze the effect of the program on academic achievement (and not only on the probability to get the "*Brevet*"), it is interesting to study pupils' scores in the written exam, since they are not affected by schools inner grading practices. Analyzing written exam

scores in French and in Mathematics give similar results as before (see Tables 1.23 and 1.24 in Appendix 1.E). For each discontinuity, there is no significant difference between pupils exogenously enrolled in RAR and pupils exogenously not enrolled in RAR, whereas a naive OLS estimation always gives a worse performance for RAR pupils.

Table 1.10 – Estimation of the effect of enrollment in a RAR on passing the *Brevet* - Discontinuity d10

	OLS			RDD (2SLS)		
	h10=1	h10=1.5	h10=2	h10=1	h10=1.5	h10=2
	Second stage					
<i>Y=Pass</i>						
<i>T=RAR</i>	-0.12***	-0.15***	-0.15***	0.01	0.05	0.23
	(0.03)	(0.03)	(0.02)	(0.07)	(0.28)	(0.34)
N	2,691	4,197	6,247	2,691	4,197	6,247
N clusters	295	460	660	295	460	660
<i>Y=Pass with honours</i>						
<i>T=RAR</i>	-0.21***	-0.21***	-0.21***	-0.07	-0.48	0.08
	(0.04)	(0.03)	(0.02)	(0.08)	(0.62)	(0.29)
N	2,691	4,197	6,247	2,691	4,197	6,247
N clusters	295	460	660	295	460	660
	First stage					
$\mathbb{1}\{Z_L > 10\}$				0.68***	0.38***	0.36***
				(0.02)	(0.05)	(0.04)
N				2,691	4,197	6,247
F-stat				835	57	74

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for different bandwidths of size h10 around each threshold.

Lecture: On average, pupils enrolled in a RAR school in 6th grade (and whose closest junior high school is near the 10% threshold) have a 12 to 15 percentage point lower probability to pass the *Brevet* than non RAR pupils. These differences are significant at the 1% level. Pupils enrolled in a RAR school exogenously, due to the fact that their closest public junior high school is above the 10% threshold, have a 1 to 23 percentage point higher probability to pass the *Brevet* than pupils exogenously not enrolled in a RAR. These differences are not significant.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.11 – Estimation of the effect of enrollment in a RAR on passing the *Brevet* - Discontinuity d67

	OLS			RDD (2SLS)		
	h67=1	h67=2	h67=3	h67=1	h67=2	h67=3
Second stage						
<i>Y=Pass</i>						
<i>T=RAR</i>	-0.10*** (0.03)	-0.16*** (0.03)	-0.15*** (0.02)	-0.07 (0.13)	0.01 (0.15)	0.21 (0.18)
N	1,402	2,441	3,679	1,402	2,441	3,679
N clusters	207	324	436	207	324	436
<i>Y=Pass with honours</i>						
<i>T=RAR</i>	-0.18*** (0.04)	-0.25*** (0.04)	-0.22*** (0.03)	-0.13 (0.18)	0.03 (0.23)	0.21 (0.24)
N	1,402	2,441	3,679	1,402	2,441	3,679
N clusters	207	324	436	207	324	436
First stage						
$\mathbb{1}\{Z_F > 67\}$				0.65*** (0.04)	0.43*** (0.03)	0.45*** (0.03)
N				1,402	2,441	3,679
F-stat				267	194	229

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for different bandwidths of size h67 around each threshold.

Lecture: On average, pupils enrolled in a RAR school in 6th grade (and whose closest junior high school is near the 67% threshold) have a 10 to 16 percentage point lower probability to pass the *Brevet* than non RAR pupils. These differences are significant at the 1% level. Pupils enrolled in a RAR school exogenously, due to the fact that their closest public junior high school is above the 67% threshold, have a probability between -7 percentage point lower to 21 percentage point higher to pass the *Brevet* than pupils exogenously not enrolled in a RAR. These differences are not significant.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Developing pupils' educational and professional aspirations was part of the objectives of the RAR program. The idea is to prevent pupils from schooling choices that would be driven by a lack of information. Thus it seems appropriate to study not only the treatment effects on pupils' academic performances, as measured by the "*Brevet*" results, but also on their situations at the end of junior high school. At the

end of 9th grade, do RAR pupils have similar or better situations than other pupils?

On average, pupils attending a RAR school in 6th grade are more often enrolled in a vocational upper secondary education track and less often in a general upper secondary education track than other pupils. However, this is mainly due to selection and sorting into and out of RAR, since RDD estimates do not show any significant treatment effect on pupils benefiting exogenously from the program (Tables 1.12 and 1.13).

The probability to repeat a grade during junior high school, and so, to be still in lower secondary education five years after entering 6th grade is almost the same for RAR and non RAR pupils in the sample. However, this probability is significantly higher for the pupils who are treated because their closest junior high school is above the 10% threshold, for the smallest bandwidth size (Table 1.12). This result is not robust to a larger bandwidth. Nonetheless, since the bias tends to increase with the bandwidth size, when the results are not robust to the size of the bandwidth, one should prefer the smallest one. This result is not observed around the 67% threshold.

Heterogeneous effects?

Whether we consider educational performance or situation after five years, enrollment in a RAR junior high school does not seem to be associated with significant academic differences, as long as pupils exogenously affected by the treatment are considered. If significant effects cannot be observed on the population of *compliers*, there may be heterogeneous effects according to individual characteristics. For instance, the program may have positive effects for some pupils and negative effects for others, resulting in a local average effect close to zero.

For that reason, it is of interest to not consider the entire population of *compliers*, but, among them, to distinguish pupils according to their observable characteristics in the data: their cohorts, their sexes, whether or not they benefit from a scholarship in 6th grade, or whether or not they come from disadvantaged social backgrounds.

For the population of *compliers* around the 67% threshold, no significant effect is detected on any sub-populations.

For the population of *compliers* around the 10% threshold, however, the treatment effect differs for different treated sub-populations. If there is no significant effect among the socially disadvantaged pupils, there seems to be negative treatment effects for more advantaged ones: their probability to pass the "*Brevet*" decreases (Table 1.14), and, in particular, they perform worse in the French written exam

Table 1.12 – Estimation of the effect of enrollment in a RAR on the situation 5 years later - Discontinuity d10

	OLS			RDD		
	h10=1	h10=1.5	h10=2	h10=1	h10=1.5	h10=2
Second stage						
<i>Y=gen high school</i>						
<i>T=RAR</i>	-0.17*** (0.03)	-0.19*** (0.02)	-0.18*** (0.02)	-0.01 (0.09)	0.42 (0.53)	-0.08 (0.21)
N	2,713	4,193	6,159	2,713	4,193	6,159
N clusters	301	461	656	301	461	656
<i>Y=pro high school</i>						
<i>T=RAR</i>	0.13*** (0.03)	0.16*** (0.02)	0.16*** (0.02)	-0.19 (0.10)	-0.69 (0.63)	-0.49 (0.50)
N	2,713	4,193	6,159	2,713	4,193	6,159
N clusters	301	461	656	301	461	656
<i>Y=late (junior high)</i>						
<i>T=RAR</i>	0.04 (0.03)	0.03 (0.02)	0.01 (0.02)	0.18*** (0.06)	0.46 (0.37)	0.46 (0.37)
N	2,713	4,193	6,159	2,713	4,193	6,159
N clusters	301	461	656	301	461	656
First stage						
$\mathbb{1}\{Z_L > 10\}$				0.68*** (0.02)	0.35*** (0.05)	0.37*** (0.04)
N				2,713	4,193	6,159
F-stat				849	45	76

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for different bandwidths of size h10 around each threshold.

Lecture: On average, pupils enrolled in a RAR school in 6th grade (and whose closest junior high school is near the 10% threshold) have a 17 to 19 percentage point lower probability than non RAR pupils to be enrolled in a general high school 5 years later. These differences are significant at the 1% level. Pupils enrolled in a RAR school exogenously, due to the fact that their closest public junior high school is above the 10% threshold, have a probability between -1 percentage point lower and 42 percentage point higher than pupils exogenously not enrolled in a RAR to be enrolled in a general high school 5 years later. These differences are not significant.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.13 – Estimation of the effect of enrollment in a RAR on the situation 5 years later - Discontinuity d67

	OLS			RDD		
	h67=1	h67=2	h67=3	h67=1	h67=2	h67=3
Second stage						
<i>Y=gen high school</i>						
<i>T=RAR</i>	-0.12**	-0.18***	-0.18***	0.03	0.16	0.24
	(0.06)	(0.05)	(0.03)	(0.16)	(0.23)	(0.22)
N	1,386	2,381	3,653	1,386	2,381	3,653
N clusters	206	320	431	206	320	431
<i>Y=pro high school</i>						
<i>T=RAR</i>	0.15***	0.17***	0.18***	0.11	0.05	-0.01
	(0.03)	(0.03)	(0.03)	(0.09)	(0.14)	(0.13)
N	1,386	2,381	3,653	1,386	2,381	3,653
N clusters	206	320	431	206	320	431
<i>Y=late (junior high)</i>						
<i>T=RAR</i>	-0.01	0.01	0.00	-0.11	-0.26	-0.21
	(0.05)	(0.05)	(0.03)	(0.11)	(0.18)	(0.13)
N	1,386	2,381	3,653	1,386	2,381	3,653
N clusters	206	320	431	206	320	431
First stage						
$\mathbb{1}\{Z_F > 67\}$				0.63***	0.41***	0.46***
				(0.04)	(0.03)	(0.03)
N				1,386	2,381	3,653
F-stat				228	162	224

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for different bandwidths of size h67 around each threshold.

Lecture: On average, pupils enrolled in a RAR school in 6th grade (and whose closest junior high school is near the 67% threshold) have a 12 to 18 percentage point lower probability than non RAR pupils to be enrolled in a general high school 5 years later. These differences are significant at the 1% level. Pupils enrolled in a RAR school exogenously, due to the fact that their closest public junior high school is above the 67% threshold, have a 3 to 24 percentage point higher probability than pupils exogenously not enrolled in a RAR to be enrolled in a general high school 5 years later. These differences are not significant.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

(Table 1.15). They are also less often enrolled in a general upper secondary education track after junior high school (Table 1.16).

Girls and boys are affected differently by the treatment, since treated girls have a higher probability to pass the "*Brevet*" than non treated ones, whereas for boys, the treatment decreases the probability to pass the exam (Table 1.14). But the average effect on the probability to repeat a grade found in Table 1.12 is not different according to individual characteristics, whatever the sub-population (Table 1.17). These differentiated effects do not hold for a larger bandwidth size (see Appendix 1.F). But, as mentioned earlier, we focus on the smallest one in this case.

From the beginning of the 2007 school year, the catchment area system was modified by the "*carte scolaire*" reform. The objectives were to give parents more freedom in their choices of schools and to reinforce social and ethnic diversity within schools. From 2007, dispensation demands were thus supposed to be made easier. Pupils entering 6th grade in 2006 were subject to the rules that prevail before the reform, whereas those entering 6th grade in 2007 were subject to a looser system. This raises two questions. First, does it constitute a confounding factor? In other words, do we attribute effects to the RAR program, which would actually be due to a change in the institutional context? Some studies showed that the reform increased the number of dispensation demands in RAR schools and thus decreased RAR schools sizes (Fack and Grenet, 2013; Murat and Thauvel-Richard, 2013). The answer is no; as soon as there is no more dispensation demands above or below the thresholds in the absence of the program (cf. Appendix 1.A), then the catchment area system reform does not constitute a confounding factor in our analysis. Secondly, because pupils were not subject to the same rules in 2006 and in 2007, then the RAR program may have had different impacts on these two cohorts. One thus need to test for differentiated treatment effects on both cohorts. The results are reported in Appendix 1.F and show that the average effects estimated on both cohorts are not significantly different from each other. This means that the average treatment effects on the population of *compliers* were not modified by the catchment area system reform.

Table 1.14 – Estimation of heterogeneous effects of enrollment in a RAR on passing the *Brevet*

Y=Pass Brevet			
	OLS	RDD d10	RDD d67
		h10=1	h67=1
<i>X=Sex (ref. = Girl)</i>			
X=0	-0.16*** (0.01)	0.18** (0.08)	-0.03 (0.13)
X=1	-0.21*** (0.01)	-0.18** (0.08)	-0.14 (0.16)
Test (pvalue)	0.000	0.000	0.436
N	1,060,069	2,691	1,402
N clusters	13,387	295	207
<i>X=Scholarship (ref. = no)</i>			
X=0	-0.14*** (0.01)	-0.07 (0.12)	-0.18 (0.15)
X=1	-0.09*** (0.01)	0.07 (0.06)	0.13 (0.12)
Test (pvalue)	0.000	0.295	0.067
N	1,060,069	2,691	1,402
N clusters	13,387	295	207
<i>X=Disadvantaged background (ref. = no)</i>			
X=0	-0.15*** (0.01)	-0.24** (0.10)	-0.10 (0.16)
X=1	-0.13*** (0.01)	0.08 (0.06)	0.02 (0.10)
Test (pvalue)	0.003	0.003	0.503
N	1,035,063	2,604	1,360
N clusters	13,376	289	206

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h10 and h67 around each threshold.

Lecture: Girls enrolled in a RAR school in 6th grade exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, have a 18 percentage point higher probability to pass the *Brevet* than girls exogenously not enrolled in a RAR. This difference is significant at the 5% level. Boys enrolled in a RAR school exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, have a 18 percentage point lower probability to pass the *Brevet* than boys exogenously not enrolled in a RAR. This difference is significant at the 5% level. The difference between these two estimates is significantly different from zero.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.15 – Estimation of heterogeneous effects of enrollment in a RAR on French grade

Y=French exam grade (/20)			
	OLS	RDD d10	RDD d67
		h10=1	h67=1
<i>X=Sex (ref. = Girl)</i>			
X=0	-1.35*** (0.05)	0.70 (0.58)	-1.45 (1.14)
X=1	-1.41*** (0.05)	-2.35*** (0.74)	-0.98 (1.28)
Test (pvalue)	0.198	0.002	0.702
N	1,051,974	2,657	1,380
N clusters	13,382	293	205
<i>X=Scholarship (ref. = no)</i>			
X=0	-1.10*** (0.06)	-3.09*** (1.05)	-1.58 (1.37)
X=1	-0.36*** (0.05)	0.68 (0.41)	0.57 (1.10)
Test (pvalue)	0.000	0.003	0.132
N	1,051,974	2,657	1,380
N clusters	13,382	293	205
<i>X=Disadvantaged background (ref. = no)</i>			
X=0	-1.16*** (0.07)	-3.19*** (0.85)	-1.36 (1.42)
X=1	-0.62*** (0.05)	0.10 (0.37)	0.24 (0.79)
Test (pvalue)	0.000	0.000	0.291
N	1,027,371	2,572	1,338
N clusters	13,372	287	204

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h10 and h67 around each threshold.

Lecture: Girls enrolled in a RAR school in 6th grade exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, get a 0.70 point (over 20) higher grade in French than girls exogenously not enrolled in a RAR. This difference is not significantly different from zero. Boys enrolled in a RAR school exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, get a 2.35 point (over 20) lower grade in French than boys exogenously not enrolled in a RAR. This difference is significant at the 1% level. The difference between these two estimates is significantly different from zero.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.16 – Estimation of heterogeneous effects of enrollment in a RAR on enrollment in general high school

Y=Gen high school			
	OLS	RDD d10	RDD d67
		h10=1	h67=1
<i>X=Sex (ref. = Girl)</i>			
X=0	-0.20*** (0.01)	0.06 (0.11)	0.06 (0.17)
X=1	-0.19*** (0.01)	-0.09 (0.11)	-0.01 (0.22)
Test (pvalue)	0.782	0.337	0.743
N	1,027,815	2,713	1,386
N clusters	13,396	301	206
<i>X=Scholarship (ref. = no)</i>			
X=0	-0.16*** (0.01)	-0.19 (0.21)	-0.10 (0.21)
X=1	-0.02*** (0.01)	0.10 (0.09)	0.34 (0.22)
Test (pvalue)	0.000	0.301	0.197
N	1,027,815	2,713	1,386
N clusters	13,396	301	206
<i>X=Disadvantaged background (ref. = no)</i>			
X=0	-0.18*** (0.01)	-0.26** (0.10)	-0.07 (0.28)
X=1	-0.05*** (0.01)	0.01 (0.07)	0.25 (0.18)
Test (pvalue)	0.000	0.009	0.357
N	1,002,986	2,620	1,344
N clusters	13,385	293	205

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h10 and h67 around each threshold.

Lecture: Girls enrolled in a RAR school in 6th grade exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, have a 6 percentage point higher probability to be enrolled in a general high school 5 years later than girls exogenously not enrolled in a RAR. This difference is not significantly different from zero. Boys enrolled in a RAR school exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, have a 9 percentage point lower probability to be enrolled in a general high school 5 years later than boys exogenously not enrolled in a RAR. This difference is not significantly different from zero. The difference between these two estimates is not significantly different from zero.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.17 – Estimation of heterogeneous effects of enrollment in a RAR on repetition

Y=Late (junior high)			
	OLS	RDD d10 h10=1	RDD d67 h67=1
<i>X=Sex (ref. = Girl)</i>			
X=0	0.02*** (0.00)	0.11 (0.06)	-0.06 (0.12)
X=1	0.01 (0.01)	0.25*** (0.08)	-0.16 (0.15)
Test (pvalue)	0.139	0.103	0.485
N	1,027,815	2,713	1,386
N clusters	13,396	301	206
<i>X=Scholarship (ref. = no)</i>			
X=0	0.01** (0.01)	0.26*** (0.10)	-0.08 (0.16)
X=1	-0.04*** (0.00)	0.14** (0.07)	-0.19 (0.14)
Test (pvalue)	0.000	0.283	0.566
N	1,027,815	2,713	1,386
N clusters	13,396	301	206
<i>X=Disadvantaged background (ref. = no)</i>			
X=0	0.02*** (0.01)	0.24*** (0.09)	0.02 (0.22)
X=1	-0.02*** (0.00)	0.17*** (0.06)	-0.23 (0.14)
Test (pvalue)	0.000	0.414	0.345
N	1,002,986	2,620	1,344
N clusters	13,385	293	205

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h10 and h67 around each threshold.

Lecture: Girls enrolled in a RAR school in 6th grade exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, have a 11 percentage point higher probability to repeat a grade during junior high school than girls exogenously not enrolled in a RAR. This difference is not significantly different from zero. Boys enrolled in a RAR school exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, have a 25 percentage point higher probability to repeat a grade during junior high school than boys exogenously not enrolled in a RAR. This difference is significant at the 1% level. The difference between these two estimates is not significantly different from zero.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

1.6 Conclusion

Our goal in this study was to identify the effect of place-based compensatory education on individual sorting across schools. To reach this goal, we have estimated the causal impact of the RAR program on school choice. Using an original geocoded data-base and a reliable identification strategy, we find that the RAR program decreases the individual probability to attend the closest public school, and increases symmetrically the probability to attend a private school, by about 30 percentage points for pupils living near a RAR school exogenously, due to the eligibility scheme. We find that the effects are heterogeneous with respect to social characteristics; they are completely driven by pupils from the most advantaged backgrounds. Once individual sorting is taken into account, we don't find any overall effect of the program on pupils' academic achievement as measured by the "*Brevet*" national exam scores and the schooling situation at the end of junior high school. The absence of mean effect may be due to heterogeneous effects in the population; for girls, enrolment in a RAR school increases the probability to pass the "*Brevet*", whereas for boys, the program decreases the probability to pass the exam, and the same probability decreases for pupils from advantaged social backgrounds.

These results may shed new light on how to evaluate place-based education policies. They make clear that it is very important to take individual sorting into account. Individuals do adjust to school-based education policies by changing schools; these adjustments are quick, are potentially of large magnitude, and are not random in the population. Not only are treated individuals different *ex ante* with respect to the general population (selection bias), but they may also select themselves into (or out of) treated schools or treated zones, resulting in a sorting bias. These findings may thus help explain some results of the literature; sorting effects may be a reason why empirical studies fail to find positive effects, or even find negative effects of compensatory education in secondary education. If more socially advantaged pupils are more likely to avoid treated schools, then the effect of the treatment is estimated on the less advantaged ones. And if less socially advantaged pupils have poorer academic performance on average, then the estimated average treatment effect is negative. This provides material for future research. First, one needs to control for individual sorting when evaluating place-based policies. Second, one may want to further investigate the part due to endogenous sorting. Third, one needs to examine the existence of dynamic peer effects.

On one hand, endogenous sorting may create adverse effects by increasing social segregation across schools. If more socially advantaged pupils (who are probably academically more advantaged on average) avoid treated schools, then those schools lose potentially "good" peers, which would be detrimental to treated pupils. But on the other hand, it could be that sorting effects lead to better target the pupils most in need of the program. Maybe those pupils are better off if only they benefit from the program. From a public policy point of view, the interpretation of our results may thus lead to a trade-off between desirable effects and adverse effects.

Our findings are true locally, for pupils who live near a school that is close to the thresholds. The findings cannot be generalized to the overall population. But they show that endogenous sorting effects do exist, and they are not negligible. Since we are comparing pupils living near schools just below and just above the eligibility thresholds, we compare *a priori* pupils in the vicinity of schools at the margin of eligibility. The schools we consider are the less disadvantaged ones, according to eligibility criteria. We could speculate that, if sorting exists in relatively less deprived areas, it could be even higher in more disadvantaged ones.

Appendix

1.A Manipulation of the forcing variables

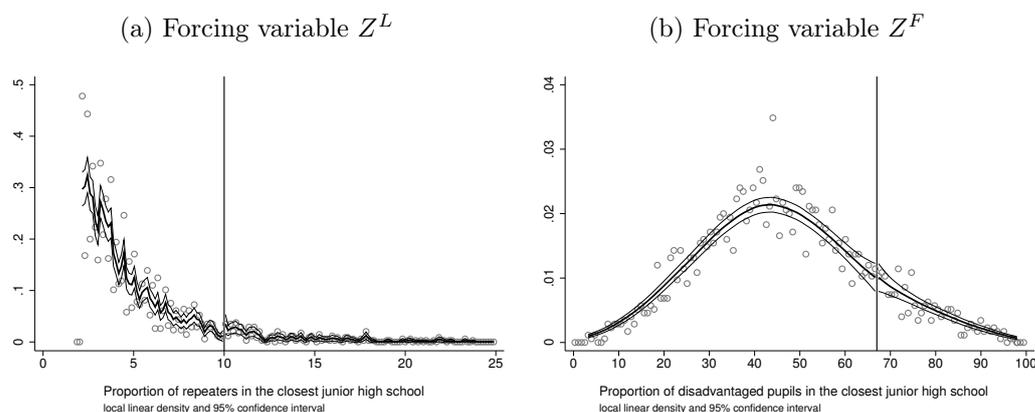
Regression discontinuity designs rely on the assumption that the forcing variable Z is continuous around the threshold ($(z^F, z^L) \mapsto \mathbb{E}(Y(j)|Z^F = z^F, Z^L = z^L)$ is continuous for $j = 0, 1$). In particular, it means that individuals cannot manipulate the forcing variable. In our case, remember that the selection variables are the proportion of repeaters in the nearest public junior high school and the proportion of pupils from disadvantaged backgrounds in the nearest public junior high school. Both variables were measured during the school year 2004-2005.

Manipulation of these variables could be the work either of junior high schools or of families living in the catchment area. In the first case, it could be that the heads of junior high schools intentionally manipulated the information relative to the number of repeaters and the number of disadvantaged pupils in their schools to fall into the eligibility group. In the second case, families could have anticipated the program and then moved in order to live closer to a school being above (or below) the thresholds.

Both scenarios are very unlikely. The first scenario would assume that the heads of junior high schools were aware in 2004 of both the measures and the cutoff values that would be chosen as eligibility criteria two years later, in 2006. The second would assume that families knew whether the nearest junior high school was below or above the thresholds and would therefore have moved closer to another school. Because the eligibility criteria and the thresholds were arbitrarily selected by the French Education Ministry so as to account for 5% of pupils, and because this information was never made public, this seems very implausible. For those reasons, we do not believe that the forcing variables may have been manipulated.

This is supported by empirical evidence. Following McCrary (2008), in case of manipulation, one would expect to find more observations above (or below) the thresholds. We thus tested for possible discontinuities in the conditional density of forcing variables. Figure 1.5 presents the local linear density estimates for both selection variables below and above the thresholds. Had headmasters misreported the proportion of pupils being late or coming from disadvantaged backgrounds in 2004 in order for their schools to enter the program, we would see many observations above the cutoff value, and few below. In the same way, had families moved closer to a school just below the threshold to be closer to a RAR school, we would see a jump at the cutoff. We do not see, however, any significant discontinuity around the thresholds.

Figure 1.5 – McCrary Density Test



Source : MEN-MESR DEPP, FAERE

1.B Placebo tests

To further test for the validity of our estimation strategy, we ran Placebo estimations, that is, we tested for the existence of discontinuities around other values than the thresholds of 10% and 67%. More precisely, we re-estimated the regression discontinuity model around Placebo values of the forcing variables.

Tables 1.18 and 1.19 present the results of estimating equations 1.1 and 1.3 respectively, around 8% and 12% of repeaters in the closest junior high school on one hand (left part of Tables), and around 57% and 77% of pupils from a disadvantaged background on the other hand (right part of Tables). In both cases, these Placebo values correspond to about one standard deviation from the true cutoff. The first stage of Table 1.18 shows that there is no significant discontinuity in the probability to live near a RAR junior high school around the 8% Placebo cutoff of repeaters in the closest junior high school, nor around the 12% cutoff. There is no significant discontinuity around the 57% Placebo cutoff of pupils from a disadvantaged background, nor around the 77% cutoff. As expected, the second stage estimations show no significant effect either. Similarly, Table 1.19 shows no significant discontinuity in the individual probability to be enrolled in a RAR school around the Placebo cutoffs, and no second stage effect either.

Table 1.18 – Placebo estimation of the effect of living near a RAR

	RD around the		RD around the	
	8% threshold	12% threshold	57% threshold	77% threshold
Second stage				
<i>Y=Enrollment in the nearest school</i>				
<i>T</i>	1.13 (2.67)	-0.36 (0.27)	4.82 (4.87)	0.38 (1.71)
<i>Y=Enrollment in another public school</i>				
<i>T</i>	0.13 (0.80)	0.87 (0.69)	-2.14 (3.12)	-4.16 (9.90)
<i>Y=Enrollment in a private school</i>				
<i>T</i>	-1.26 (2.57)	-0.51 (0.49)	-2.67 (2.82)	3.78 (8.73)
First stage				
$\mathbb{1}\{Z^j > c^j\}$	0.07 (0.15)	0.31 (0.23)	-0.06 (0.06)	-0.08 (0.19)
F-stat	30	377	40	9
Nbr obs	22,456	8,861	2,752	3,083
Nbr clusters	216	90	22	30

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the closest junior high school level. Two-stage least squares are estimated for a bandwidth of size $h=2$ around the threshold.

Lecture: Pupils living in the vicinity of a junior high school that is just above the 8% threshold of repeaters have a 0.07 higher probability that their closest junior high school is a RAR, but this coefficient is not significantly different from zero.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.19 – Placebo estimation of the effect of enrollment in a RAR

	RD around the		RD around the	
	8% threshold	12% threshold	57% threshold	77% threshold
Second stage				
<i>Y=Pass Brevet</i>				
<i>T</i>	-0.50 (0.37)	0.08 (0.31)	1.31 (4.21)	0.39 (1.42)
<i>Y=Pass with honours</i>				
<i>T</i>	-0.67 (0.51)	-0.25 (0.39)	0.60 (2.98)	-0.60 (1.01)
First stage				
$\mathbb{1}\{Z^j > c^j\}$	0.09 (0.07)	0.12 (0.11)	-0.01 (0.07)	0.12 (0.24)
F-stat	65	38	1	10
Nbr obs	21,135	8,408	2,583	2,928
Nbr clusters	1,838	867	335	373

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the junior high school level. Two-stage least squares are estimated for a bandwidth of size $h=2$ around the threshold.

Lecture: Pupils living in the vicinity of a junior high school that is just above the 8% threshold of repeaters have a 0.09 higher probability to be enrolled in a RAR in 6th grade, but this coefficient is not significantly different from zero.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

1.C Mean comparison of pupils' municipality characteristics

Table 1.20 – Characteristics of pupils’ municipality of residence around discontinuity d^L

	Mean comparison of:						
	Pupils living near a RAR vs not				Pupils above vs below d^L		
	Total	h=0.5	h=1	h=2	h=0.5	h=1	h=2
Median households revenue	-4512*** (362.7)	-2238* (1189.9)	-2382** (1138.8)	-1372* (807.9)	-2912 (2750.4)	-1222 (2849.1)	-1814 (2055.1)
Unemployment rate	1.04*** (0.19)	1.13** (0.54)	0.56 (0.37)	0.70** (0.28)	1.80 (1.31)	-0.75 (1.04)	0.40 (0.92)
Population density	1962*** (292.9)	1298 (904.9)	1162 (993.8)	891 (584.3)	3043** (1536.2)	-1401 (1877.4)	282 (1341.3)
Nbr obs	1,006,801	1,299	5,256	13,801	1,299	5,256	13,801

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the municipality (of residence) level. Mean differences are estimated for different bandwidth of size h_L around the 10% threshold.

Lecture: Pupils living in the vicinity of a RAR junior high school live in municipalities where the median households revenue is 4512 euros smaller than in other pupils’ municipalities. This difference is significant at the 1% level and remains significant for different sub-samples of observations lying just around the 10% threshold.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.21 – Characteristics of pupils’ municipality of residence around discontinuity d^F

	Mean comparison of:						
	Pupils living near a RAR vs not				Pupils above vs below d^F		
	Total	h=2	h=4	h=8	h=2	h=4	h=8
Median households revenue	-4512*** (362.7)	1155 (1282.2)	-2022 (1241.9)	-1027 (752.5)	1259 (2794.8)	2780 (2046.4)	-305 (1634.2)
Unemployment rate	1.04*** (0.19)	0.23 (0.64)	0.03 (0.36)	0.23 (0.21)	0.49 (1.13)	-0.73 (0.89)	-0.83* (0.49)
Population density	1962*** (292.9)	1432 (1259.4)	788 (878.1)	-109 (588.6)	3317 (2105.7)	688 (1818.8)	-1614 (1314.1)
Nbr obs	1,006,801	2,496	5,039	14,367	2,496	5,039	14,367

55

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the municipality (of residence) level. Mean differences are estimated for different bandwidth of size h_F around the 67% threshold.

Lecture: Pupils living in the vicinity of a RAR junior high school live in municipalities where the population density is 1962 inhabitants per square kilometer higher than in other pupils’ municipalities. This difference is significant at the 1% level and turns not significant in the sub-samples of observations lying just around the 67% threshold.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

1.D Results around discontinuity d^F

Table 1.22 – Estimation of heterogeneous effects of living near a RAR on school choice around discontinuity d^F

Y=Enrollment in	the nearest school			another public school			a private school		
	RD around the 67% threshold								
	h=2	h=4	h=8	h=2	h=4	h=8	h=2	h=4	h=8
<i>X=Scholarship (ref.=no)</i>									
<i>X=0</i>	0.19 (0.29)	-0.25 (0.18)	-0.16 (0.20)	-0.09 (0.27)	-0.19 (0.23)	0.03 (0.23)	-0.10 (0.15)	0.44* (0.23)	0.13 (0.25)
<i>X=1</i>	-0.12 (0.27)	-0.18 (0.18)	-0.24 (0.18)	0.12 (0.27)	0.03 (0.14)	0.13 (0.17)	0.00 (0.08)	0.15 (0.09)	0.11* (0.06)
Test (pvalue)	0.426	0.620	0.721	0.605	0.394	0.697	0.425	0.111	0.924
Nbr obs	2,548	5,602	15,022	2,548	5,602	15,022	2,548	5,602	15,022
Nbr clusters	30	58	144	30	58	144	30	58	144
<i>X=Disadvantaged background (ref.=no)</i>									
<i>X=0</i>	0.10 (0.24)	-0.37** (0.18)	-0.31 (0.19)	-0.12 (0.27)	-0.15 (0.23)	0.15 (0.20)	0.02 (0.17)	0.53* (0.27)	0.15 (0.23)
<i>X=1</i>	-0.06 (0.19)	-0.19 (0.18)	-0.10 (0.16)	0.16 (0.20)	0.00 (0.13)	0.01 (0.17)	-0.10 (0.09)	0.18* (0.10)	0.09 (0.09)
Test (pvalue)	0.454	0.177	0.254	0.371	0.485	0.500	0.467	0.094	0.720
Nbr obs	2,475	5,416	14,392	2,475	5,416	14,392	2,475	5,416	14,392
Nbr clusters	30	58	144	30	58	144	30	58	144

57

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the closest junior high school level. Two stage least squares are estimated for different bandwidths of size h around the threshold.

Lecture: Scholarship pupils who live near a RAR junior high school exogenously, due to the fact that their closest public junior high school is just above the 67% threshold ($h=0.5$), have a 17 percentage point smaller probability to enroll in another public junior high school than scholarship pupils whose nearest junior high school is not a RAR exogenously. Non scholarship pupils who live near a RAR junior high school exogenously, due to the fact that their closest public junior high school is just above the 10% threshold, have a 29 percentage point smaller probability to enroll in another public junior high school than non scholarship pupils whose nearest junior high school is not a RAR exogenously. The difference between these two estimates is not significantly different from zero.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

1.E Estimations on grades

Table 1.23 – Estimation of the effect of enrollment in a RAR on *Brevet* grades - Discontinuity d10

	OLS			RDD (2SLS)		
	h10=1	h10=1.5	h10=2	h10=1	h10=1.5	h10=2
Second stage						
<i>Y=Exam grade (/20)</i>						
<i>T=RAR</i>	-1.95***	-2.11***	-2.09***	0.10	-2.23	1.87
	(0.39)	(0.28)	(0.21)	(0.87)	(4.79)	(3.56)
N	2,654	4,144	6,171	2,654	4,144	6,171
N clusters	293	457	656	293	457	656
<i>Y=French exam grade (/20)</i>						
<i>T=RAR</i>	-0.68**	-0.71***	-0.68***	-0.76	-4.99	-0.55
	(0.27)	(0.23)	(0.19)	(0.45)	(4.72)	(1.61)
N	2,654	4,144	6,171	2,654	4,144	6,171
N clusters	293	457	656	293	457	656
<i>Y=Math exam grade (/20)</i>						
<i>T=RAR</i>	-0.89***	-1.13***	-1.22***	0.27	0.23	0.44
	(0.31)	(0.23)	(0.19)	(0.82)	(3.86)	(2.32)
N	2,654	4,144	6,171	2,654	4,144	6,171
N clusters	293	457	656	293	457	656
<i>Y=Brevet average (/20)</i>						
<i>T=RAR</i>	-1.29***	-1.40***	-1.42***	-0.24	-2.27	0.60
	(0.28)	(0.22)	(0.17)	(0.65)	(3.76)	(2.21)
N	2,654	4,144	6,171	2,654	4,144	6,171
N clusters	293	457	656	293	457	656
<i>Y=CC average (/20)</i>						
<i>T=RAR</i>	-0.87***	-0.95***	-1.00***	-0.41	-2.09	-0.05
	(0.24)	(0.20)	(0.17)	(0.60)	(3.36)	(1.66)
N	2,654	4,144	6,171	2,654	4,144	6,171
N clusters	293	457	656	293	457	656
First stage						
$\mathbb{1}\{Z_L > 10\}$				0.68***	0.37***	0.35***
				(0.02)	(0.05)	(0.04)
N				2,654	4,144	6,171
F-stat				825	54	71

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for different bandwidths of size h10 around each threshold.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.24 – Estimation of the effect of enrollment in a RAR on *Brevet* grades - Discontinuity d67

	OLS			RDD (2SLS)		
	h67=1	h67=2	h67=3	h67=1	h67=2	h67=3
	Second stage					
<i>Y=Exam grade (/20)</i>						
<i>T=RAR</i>	-1.93*** (0.34)	-2.48*** (0.30)	-2.13*** (0.28)	-1.66 (1.75)	1.03 (2.44)	2.22 (2.40)
N	1,379	2,409	3,634	1,379	2,409	3,634
N clusters	204	321	432	204	321	432
<i>Y=French exam grade (/20)</i>						
<i>T=RAR</i>	-1.15*** (0.25)	-1.38*** (0.25)	-1.43*** (0.19)	-0.92 (1.09)	-0.82 (1.19)	0.75 (1.30)
N	1,379	2,409	3,634	1,379	2,409	3,634
N clusters	204	321	432	204	321	432
<i>Y=Math exam grade (/20)</i>						
<i>T=RAR</i>	-1.26*** (0.38)	-1.88*** (0.33)	-1.76*** (0.22)	-0.85 (1.56)	-0.30 (1.81)	1.25 (1.75)
N	1,379	2,409	3,634	1,379	2,409	3,634
N clusters	204	321	432	204	321	432
<i>Y=Brevet average (/20)</i>						
<i>T=RAR</i>	-1.33*** (0.28)	-1.77*** (0.24)	-1.58*** (0.20)	-1.26 (1.29)	0.05 (1.59)	1.53 (1.70)
N	1,379	2,409	3,634	1,379	2,409	3,634
N clusters	204	321	432	204	321	432
<i>Y=CC average (/20)</i>						
<i>T=RAR</i>	-0.92*** (0.24)	-1.29*** (0.21)	-1.22*** (0.16)	-1.09 (0.99)	-0.68 (1.10)	1.04 (1.24)
N	1,379	2,409	3,634	1,379	2,409	3,634
N clusters	204	321	432	204	321	432
	First stage					
$\mathbb{1}\{Z_F > 67\}$				0.65*** (0.04)	0.43*** (0.03)	0.45*** (0.03)
N				1,379	2,409	3,634
F-stat				259	187	223

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for different bandwidths of size h67 around each threshold.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

1.F Heterogeneous effects on achievement

Table 1.25 – Estimation of heterogeneous effects of enrollment in a RAR on passing the *Brevet*

	Y=Pass Brevet						
	OLS	RDD d10			RDD d67		
		h10=1	h10=1.5	h10=2	h67=1	h67=2	h67=3
<i>X=Cohort (ref. = 2006)</i>							
X=0	-0.19*** (0.01)	0.00 (0.10)	0.07 (0.29)	0.34 (0.64)	-0.02 (0.17)	0.29 (0.39)	0.35 (0.33)
X=1	-0.18*** (0.01)	0.01 (0.10)	-0.03 (0.63)	0.13 (0.36)	-0.09 (0.17)	-0.13 (0.17)	0.12 (0.20)
Test (pvalue)	0.127	0.984	0.889	0.775	0.756	0.313	0.559
N	1,060,069	2,691	4,197	6,247	1,402	2,441	3,679
N clusters	13,387	295	460	660	207	324	436
<i>X=Sex (ref. = Girl)</i>							
X=0	-0.16*** (0.01)	0.18** (0.08)	0.51 (0.39)	0.26 (0.36)	-0.03 (0.13)	0.07 (0.16)	0.23 (0.17)
X=1	-0.21*** (0.01)	-0.18** (0.08)	-0.48 (0.49)	0.18 (0.39)	-0.14 (0.16)	-0.07 (0.17)	0.18 (0.20)
Test (pvalue)	0.000	0.000	0.143	0.786	0.436	0.341	0.675
N	1,060,069	2,691	4,197	6,247	1,402	2,441	3,679
N clusters	13,387	295	460	660	207	324	436
<i>X=Scholarship (ref. = no)</i>							
X=0	-0.14*** (0.01)	-0.07 (0.12)	-0.76 (1.01)	-0.03 (0.31)	-0.18 (0.15)	0.01 (0.20)	0.21 (0.20)
X=1	-0.09*** (0.01)	0.07 (0.06)	0.37 (0.36)	0.31 (0.28)	0.13 (0.12)	0.04 (0.13)	0.08 (0.07)
Test (pvalue)	0.000	0.295	0.372	0.342	0.067	0.905	0.499
N	1,060,069	2,691	4,197	6,247	1,402	2,441	3,679
N clusters	13,387	295	460	660	207	324	436
<i>X=Disadvantaged background (ref. = no)</i>							
X=0	-0.15*** (0.01)	-0.24** (0.10)	-0.88 (0.79)	-0.09 (0.22)	-0.10 (0.16)	-0.04 (0.21)	0.32 (0.29)
X=1	-0.13*** (0.01)	0.08 (0.06)	0.24 (0.33)	0.57 (0.75)	0.02 (0.10)	0.13 (0.14)	0.16 (0.10)
Test (pvalue)	0.003	0.003	0.290	0.368	0.503	0.431	0.517
N	1,035,063	2,604	4,064	6,037	1,360	2,375	3,578
N clusters	13,376	289	448	644	206	321	429

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h10 and h67 around each threshold.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.26 – Estimation of heterogeneous effects of enrollment in a RAR on Math grade

	Y=Math exam grade (/20)						
	OLS	RDD d10			RDD d67		
		h10=1	h10=1.5	h10=2	h67=1	h67=2	h67=3
<i>X=Cohort (ref. = 2006)</i>							
X=0	-1.79*** (0.08)	0.12 (1.12)	-1.39 (4.25)	2.23 (4.76)	-0.23 (2.06)	2.41 (4.02)	3.20 (3.41)
X=1	-1.73*** (0.08)	0.51 (0.93)	2.84 (7.08)	-1.20 (3.01)	-1.03 (2.06)	-1.63 (2.18)	-0.02 (1.95)
Test (pvalue)	0.593	0.786	0.609	0.543	0.784	0.377	0.413
N	1,051,790	2,656	4,146	6,174	1,380	2,410	3,638
N clusters	13,382	293	457	656	204	321	432
<i>X=Sex (ref. = Girl)</i>							
X=0	-1.67*** (0.06)	1.35 (0.81)	3.21 (3.28)	1.95 (2.97)	-0.76 (1.58)	0.35 (2.45)	1.47 (2.08)
X=1	-1.87*** (0.06)	-0.96 (1.12)	-3.35 (7.44)	-1.07 (2.78)	-1.24 (2.01)	-1.06 (1.96)	1.01 (1.83)
Test (pvalue)	0.001	0.022	0.414	0.361	0.809	0.585	0.784
N	1,051,790	2,656	4,146	6,174	1,380	2,410	3,638
N clusters	13,382	293	457	656	204	321	432
<i>X=Scholarship (ref. = no)</i>							
X=0	-1.46*** (0.08)	-1.77 (1.40)	-10.46 (14.03)	-4.04 (4.08)	-1.70 (2.09)	0.54 (2.68)	2.34 (2.54)
X=1	-0.39*** (0.05)	1.63 (0.90)	4.84 (4.20)	2.51 (1.91)	1.23 (1.29)	-1.01 (1.43)	-0.47 (0.73)
Test (pvalue)	0.000	0.070	0.365	0.173	0.184	0.566	0.271
N	1,051,790	2,656	4,146	6,174	1,380	2,410	3,638
N clusters	13,382	293	457	656	204	321	432
<i>X=Disadvantaged background (ref. = no)</i>							
X=0	-1.64*** (0.08)	-2.08 (1.07)	-13.14 (11.15)	-1.94 (2.70)	-1.70 (2.01)	0.54 (3.22)	3.27 (3.66)
X=1	-0.77*** (0.05)	0.96 (0.74)	5.65 (6.45)	3.87 (4.64)	0.69 (1.20)	0.18 (1.36)	0.33 (0.83)
Test (pvalue)	0.000	0.005	0.256	0.284	0.262	0.910	0.400
N	1,027,213	2,571	4,018	5,971	1,338	2,344	3,537
N clusters	13,372	287	445	640	203	318	425

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h10 and h67 around each threshold.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.27 – Estimation of heterogeneous effects of enrollment in a RAR on French grade

Y=French exam grade (/20)							
	OLS	RDD d10			RDD d67		
		h10=1	h10=1.5	h10=2	h67=1	h67=2	h67=3
<i>X=Cohort (ref. = 2006)</i>							
X=0	-1.38*** (0.06)	-0.43 (0.63)	-4.31 (4.70)	-0.62 (2.55)	-2.01 (1.76)	0.36 (2.26)	1.92 (2.39)
X=1	-1.33*** (0.06)	-1.16 (0.68)	-5.94 (10.55)	-0.40 (1.97)	-0.22 (1.34)	-1.40 (1.49)	-0.01 (1.50)
Test (pvalue)	0.617	0.434	0.888	0.945	0.420	0.516	0.493
N	1,051,974	2,657	4,148	6,175	1,380	2,410	3,637
N clusters	13,382	293	457	656	205	322	433
<i>X=Sex (ref. = Girl)</i>							
X=0	-1.35*** (0.05)	0.70 (0.58)	-0.58 (1.93)	1.25 (2.46)	-1.45 (1.14)	-0.45 (1.44)	0.91 (1.44)
X=1	-1.41*** (0.05)	-2.35*** (0.74)	-9.76 (8.84)	-2.45 (2.09)	-0.98 (1.28)	-1.41 (1.34)	0.40 (1.41)
Test (pvalue)	0.198	0.002	0.267	0.226	0.702	0.523	0.643
N	1,051,974	2,657	4,148	6,175	1,380	2,410	3,637
N clusters	13,382	293	457	656	205	322	433
<i>X=Scholarship (ref. = no)</i>							
X=0	-1.10*** (0.06)	-3.09*** (1.05)	-18.57 (17.49)	-4.43 (3.97)	-1.58 (1.37)	-0.63 (1.59)	0.97 (1.62)
X=1	-0.36*** (0.05)	0.68 (0.41)	1.70 (1.97)	1.66 (1.53)	0.57 (1.10)	-0.93 (1.19)	-0.18 (0.57)
Test (pvalue)	0.000	0.003	0.280	0.197	0.132	0.851	0.452
N	1,051,974	2,657	4,148	6,175	1,380	2,410	3,637
N clusters	13,382	293	457	656	205	322	433
<i>X=Disadvantaged background (ref. = no)</i>							
X=0	-1.16*** (0.07)	-3.19*** (0.85)	-15.18 (10.46)	-2.69 (2.43)	-1.36 (1.42)	-0.97 (1.82)	1.68 (2.35)
X=1	-0.62*** (0.05)	0.10 (0.37)	-0.79 (1.63)	2.03 (3.39)	0.24 (0.79)	0.07 (0.94)	0.35 (0.61)
Test (pvalue)	0.000	0.000	0.180	0.275	0.291	0.571	0.546
N	1,027,371	2,572	4,020	5,972	1,338	2,344	3,536
N clusters	13,372	287	445	640	204	319	426

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h10 and h67 around each threshold.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Table 1.28 – Estimation of heterogeneous effects of enrollment in a RAR on enrollment in general high school

Y=Gen high school							
	OLS	RDD d10			RDD d67		
		h10=1	h10=1.5	h10=2	h67=1	h67=2	h67=3
<i>X=Cohort (ref. = 2006)</i>							
X=0	-0.20*** (0.01)	-0.00 (0.13)	0.27 (0.44)	-0.01 (0.35)	0.33 (0.37)	0.61 (0.73)	0.50 (0.48)
X=1	-0.19*** (0.01)	-0.02 (0.12)	0.67 (1.55)	-0.13 (0.26)	-0.11 (0.18)	-0.07 (0.21)	0.06 (0.21)
Test (pvalue)	0.456	0.931	0.801	0.786	0.283	0.369	0.403
N	1,027,815	2,713	4,193	6,159	1,386	2,381	3,653
N clusters	13,396	301	461	656	206	320	431
<i>X=Sex (ref. = Girl)</i>							
X=0	-0.20*** (0.01)	0.06 (0.11)	0.38 (0.53)	0.06 (0.30)	0.06 (0.17)	0.39 (0.37)	0.33 (0.26)
X=1	-0.19*** (0.01)	-0.09 (0.11)	0.46 (0.72)	-0.23 (0.27)	-0.01 (0.22)	-0.07 (0.22)	0.13 (0.21)
Test (pvalue)	0.782	0.337	0.913	0.415	0.743	0.193	0.271
N	1,027,815	2,713	4,193	6,159	1,386	2,381	3,653
N clusters	13,396	301	461	656	206	320	431
<i>X=Scholarship (ref. = no)</i>							
X=0	-0.16*** (0.01)	-0.19 (0.21)	-0.25 (0.82)	-0.66 (0.48)	-0.10 (0.21)	0.07 (0.25)	0.25 (0.22)
X=1	-0.02*** (0.01)	0.10 (0.09)	0.56 (0.71)	0.27 (0.27)	0.34 (0.22)	0.24 (0.23)	0.08 (0.11)
Test (pvalue)	0.000	0.301	0.541	0.163	0.197	0.611	0.499
N	1,027,815	2,713	4,193	6,159	1,386	2,381	3,653
N clusters	13,396	301	461	656	206	320	431
<i>X=Disadvantaged background (ref. = no)</i>							
X=0	-0.18*** (0.01)	-0.26** (0.10)	-0.92 (0.67)	-0.32 (0.23)	-0.07 (0.28)	0.34 (0.46)	0.41 (0.42)
X=1	-0.05*** (0.01)	0.01 (0.07)	0.72 (0.98)	0.19 (0.36)	0.25 (0.18)	0.20 (0.20)	0.14 (0.12)
Test (pvalue)	0.000	0.009	0.292	0.275	0.357	0.744	0.510
N	1,002,986	2,620	4,054	5,940	1,344	2,312	3,549
N clusters	13,385	293	448	639	205	317	425

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Standard errors in brackets are clustered at the attended junior high school level. Two-stage least squares are estimated for bandwidths of size h10 and h67 around each threshold.

Source: MEN-MESR DEPP, FAERE 2006 and 2007

Chapter 2

Geographical constraints in track choices: a French study using high school openings

This chapter is based on joint work with Son Thierry LY and Meryam ZAIEM

Introduction

Both in the field and in the literature, there has been an increasing interest in school choice reforms over the recent years. Those reforms aim at giving parents more freedom in the choice of school for their child(ren). As an example, since 2001 in the United States, parents have been allowed to choose a school different from their catchment area one if it fails to improve pupils' results for two consecutive years (No Child Left Behind Act). In France, the catchment area system has been relaxed since 2007 allowing parents to choose more freely the middle and high school of their children.

Such school choice reforms are expected to improve the performance of the educational system, by creating demand-side pressure on under-performing schools and by better matching schools and pupils. For these reforms to be effective, one need to assume first that households are constrained in their school choice and second that they know better which school is the best for their child (Hastings, Kane, and Staiger, 2009). In this paper, we try to assess the validity of the first assumption in a specific context, namely track choice at the end of lower secondary education in France. Does pupils' allocation change when the local supply of schooling is increased by the opening of a new high school? What can we tell about the supply constraints pupils face in their track choices?

These questions relate to two strands of the literature. A first strand of the

literature studies the effect of school choice reforms. In the French context, Fack and Grenet (2012) showed that the catchment area system reform had no effect on school choice in the sense that it did not significantly increase the number of pupils asking for another school than their catchment area one. In the United States, although the context is very different, Hastings, Kane, and Staiger (2009) find that a school choice plan in North Carolina had a significant impact on school choice but ambiguous effects on academic outcomes, and Deming et al. (2014) find that attending a first-choice school increases college attainment. Studying a Tel-Aviv school choice program, Lavy (2010) shows that choice reduces the drop-out rate and increases high school achievement. The author also finds long-term positive effects on post-secondary enrollment and earnings (Lavy, 2015). A second strand of the literature shows that school supply is a major determinant of schooling rate, academic achievement and schooling decisions. A first group of papers studies the link between schooling supply and enrollment rate. They show that the probability of going to school increases significantly when new schools are built and when the distance to school decreases (Burde and Linden, 2013; Duflo, 2001; Filmer, 2007; Handa, 2002). A second group of papers focus on the link between school accessibility and pupils' performance. They identify a negative impact of distance to school on academic achievement (Burde and Linden, 2013; Falch, Lujala, and Strøm, 2013). Finally, a third group of papers points out that the local school supply is key to explain whether or not individuals pursue in higher education (Dickerson and McIntosh, 2013; Frenette, 2009; Gibbons and Vignoles, 2012; Griffith and Rothstein, 2009; Spiess and Wrohlich, 2010).

Nevertheless, the relationship between school supply and schooling decisions is complex and isolating the impact of the former on the latter is not an easy task. In fact, pupils are not located randomly relative to schools. First, schools are not evenly distributed on the territory. In France, at the beginning of the 2013 school year, there was an average of 8 high schools for every 10 000 pupils enrolled in secondary schooling. There were 6 high schools for 10 000 pupils in the Versailles school district and 13 for 10 000 pupils in the Paris district. Second, households pay attention to the school supply in the neighbourhood when choosing a house (Barrow, 2002; Bayer, Ferreira, and McMillan, 2007; Chumacero, Gómez, and Paredes, 2011; Epple and Romano, 2003; Fack and Grenet, 2010). Unobserved characteristics of households may explain both their location (and thus the school supply they face) and their track choices. Assume for example that parents with high preferences

for academic achievement are those who locate in neighbourhoods where the school supply is abundant and also those with children who study the longest. Then the quantity of local school supply is not exogenous from schooling preferences.

Our contribution to the literature is twofold. First, the literature on school supply and schooling decisions focuses mainly on primary education, or higher education. Little is known about schooling decisions at the secondary level. We do think that looking at choices at the end of middle school is interesting, especially in the case of France, where pupils make an important choice at the end of 9th grade between vocational and general tracks. This choice has long run consequences on both achieved level of schooling and labor market outcomes. Goux, Gurgand, and Maurin (2015) show that getting more low achieving pupils to follow a vocational track after middle school leads to a significant and important reduction in grade repetition and high-school drop out for those at-the-margin students. Second, we rely on high school openings to highlight the constraint local school supply exerts on pupils' schooling decisions. We use exhaustive data on 9th grade pupils from 2002 to 2011. As we are able to precisely locate middle and high schools, we are able to observe whether a high school opened in the neighbourhood of a given middle school a given year. A generalized difference in differences estimation allows us to make use of the variation in time and space of high school openings to identify the causal effect of a change in local school supply on the allocation of pupils at the end of middle school. Our results show that opening a new high school significantly impacts the allocation at the end of 9th grade for pupils from treated middle schools. The constrained pupils seem to be pupils who would like to follow a vocational track.

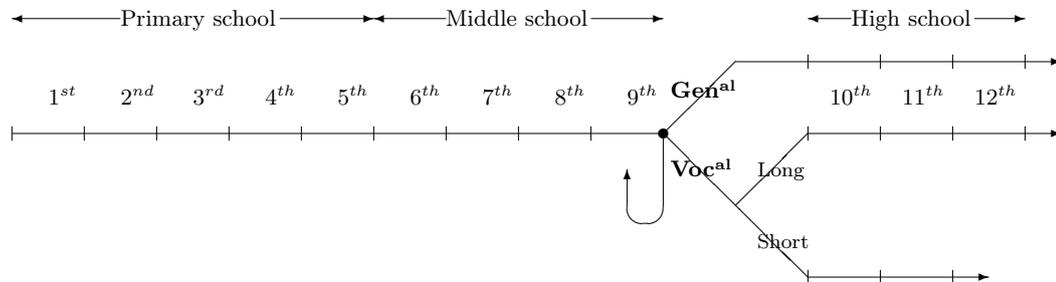
The paper is organized as follows. We first describe the institutional context of track choices at the end of 9th grade and the administrative process of opening a new high school. We then describe the data and the estimation strategy. Another section presents some descriptive statistics. Estimation results come in the last section and we conclude with a discussion.

2.1 Institutional Context

2.1.1 Track choice at the end of 9th grade

Education is compulsory in France from age 6 to age 16. Primary education lasts 5 years (from age 6 to age 10). Secondary education is divided between 4 years of lower secondary (from age 11 to age 14) in *collèges*, which are equivalent to middle

Figure 2.1 – Illustration of the French school system



schools, and 3 years of upper secondary (from age 15 to age 17) in *lycées*, equivalent to high schools.

At the end of middle school, pupils have to choose whether they would like to continue in a general or a vocational track (see Figure 2.1). In the general track, pupils study academic and technical subjects during three years to prepare for a general national exam (called *Baccalauréat*). The general *Baccalauréat* gives access to higher education. The vocational track provides a professional training. There are two types of vocational track. A two year track prepares for a professional certificate and a direct entry into the labor market as a skilled worker. A three year track prepares for a vocational *Baccalauréat* which gives access to highly qualified professions or to higher education.¹ At the end of middle school, pupils may also choose to drop out, or to repeat 9th grade if their results are not sufficient to continue in high school.

The track choice procedure starts in January of 9th grade. Families have to choose between general track, vocational track, or repetition. At the end of the second term, the teaching staff responds to families' choices by providing temporary propositions of allocation. Before the end of the third and last term, families are asked to make a final choice. If their choice matches the school recommendation, the pupil is officially allocated to this track. If the school and the family disagree on the allocation, a meeting with the school headmaster is organised. If no agreement is reached at the end of the meeting, the family may resort to an appeal board. The

¹Both vocational tracks may be completed through a work-based training (apprenticeship), which we do not study here.

decision of this board is final. However, whatever the decision, pupils are always free to choose to repeat 9th grade and go through the process again the next year.

After a decision about the track is made, pupils are allocated to high schools on the basis of a catchment area system. Each pupil has priority in the public general high school of her district according to where she lives. Pupils can go to another public high school through a special dispensation. If the number of dispensations exceeds the number of places in a given high school, the priority order is determined by the local education authority (*académie*) director. Allocation to a vocational high school is not based on the catchment area system, but on pupils' academic achievement. Another option is to go to a private high school, which is not subject to the catchment area system either.

2.1.2 Building new high schools

In France, the State and the three local authorities (*régions*, *départements*, and municipalities) share the responsibility for education. The State is responsible for defining the national curricula, delivering degrees, recruiting and paying teachers. *Régions* are responsible for high schools, meaning that they are responsible for the building, maintaining and functioning of high schools. *Départements* are responsible for middle schools. And municipalities are responsible for primary schools.

Deciding to build a new high school is a long process reflecting the sharing of responsibilities between these different entities. First, on a regular basis, regions have to plan their needs in terms of middle schools and high schools, based on the demographic situation in the region and the expected number of future pupils. Second, the representative of the State at the region level approves of the region's project. If a high school needs to be built, the regional assembly then votes to allow the building. The whole building process (from selecting a service provider to realization) often takes many years. The mean duration between the regional assembly vote and the delivery of a new high school is 5 years and the cost is between 20 and 60 million euros.

The process may be slightly different in the private sector since anyone can open a new private high school, though with prior notification to the local education officer (*recteur d'académie*). However, in France, almost all private schools are publicly-funded. They follow exactly the same national curriculum as public schools (except for religious education²) and prepare for the same national exams, their teachers are

²Most private schools (more than 95%) are Catholic schools.

employed by the State and local authorities are in charge of their functioning, in the exact same way as for public schools. About 20% of secondary education pupils are enrolled in a private school. 98% of them go to a publicly-funded school.

2.2 Data

We use exhaustive individual-level data from the SCOLARITÉ data sets. These annual databases are provided by the statistical service of the French ministry of Education and are compounded of every pupil enrolled in a secondary school supervised by the ministry of Education every year.³ In every cohort from 2002-2003 to 2011-2012 we focus on the 9th grade pupils enrolled in a middle school in mainland France (i.e. metropolitan France with the exception of Corsica). Each of these pupils are observed in year t (the year of their 9th grade) and year $t+1$. The data provide the school and track of each pupil in years t and $t+1$. We know whether they are enrolled in a private or in a public school, whether it is a middle school or a high school, and whether it is a general or a vocational high school. For each pupil, we observe some socio-demographic characteristics: age, gender, citizenship and the occupation of the parents.

A second source of data comes from an exhaustive school-level panel data set, which provides information on every French school. We focus on mainland France middle schools and high schools. Their postal address is known, so that we can observe their exact geographic location. The exact administrative date when they opened (and, if they ceased to exist, the date when they closed) is also observed.

Working with exhaustive data sets, we are able to identify, every year, high schools that appear for the first time in the data. For a given year t , a high school is considered as a new high school if some pupils are enrolled in that high school in year t while no pupil were enrolled in there the previous years. We also check that this year corresponds to the administrative date of opening. Thus we can only observe openings for cohorts starting from 2003 (the second year of observation). Moreover, a high school that appears only one year in the data set is not considered as an opening. As consequence, the last cohort of the data (2011) is excluded from the working sample because we cannot know if the openings observed that year are permanent or not.

³Since the ministry of Education is not responsible for apprenticeship, we do not observe apprentices in the data set.

A pupil is then considered as treated if a high school opened in her middle school's neighbourhood the year of her 9th grade. The treatment is thus defined at the middle school level. We tried different definitions of whether a middle school is treated or not. First, only the closest middle school to each opening high school is considered as treated. Then treated schools are extended to the two closest - and the five closest - schools to each opening high school. Second, we used an alternative definition in which treated middle schools are those which neighborhood contains an opening high school. The neighbourhood of a middle school is defined as the circle of radius r centered in the middle school, where r is equal to the median distance between the middle school and all high schools, weighted by the proportion of pupils going to each high school.

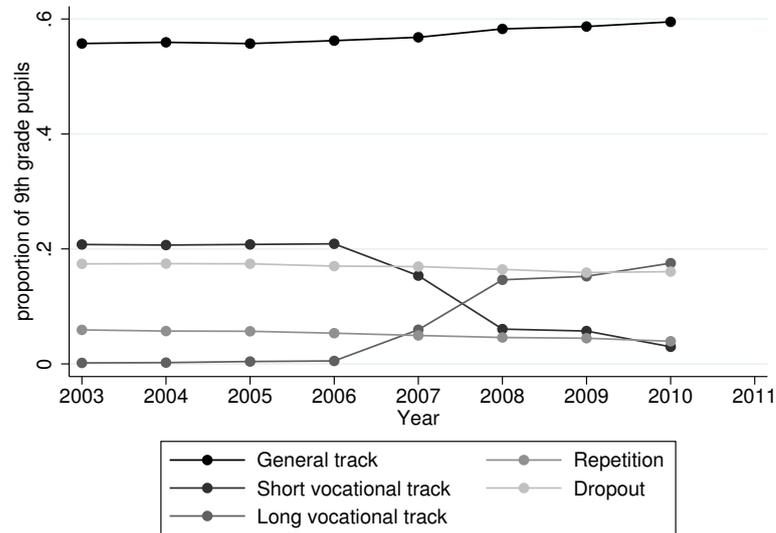
All these treatments are computed separately for different types of high schools. In France there are three types of high schools; those preparing for general tracks (*lycées généraux et technologiques* or LGT), those offering vocational tracks (*lycées professionnels* or LPR), and those providing both general and vocational tracks (called *lycées polyvalents*, hereafter LPO). Vocational high schools are less numerous and have a larger area of influence. Pupils going to a vocational high school have an average distance from middle school to high school of about 16 km, compared to 9 km for pupils attending a general high school. Thus we need to compute separate distances and treatments.

These definitions may be ranked from the more conservative (i.e. only the closest school is treated) to the less conservative (i.e. all schools with a new high school within their radius are treated). According to the first definition, there are as much treated middle schools as opening high schools; with the second definition, there are two treated schools for every new high school; with the third definition, there are five treated schools for every new high school; in the last case, there are about 15 treated schools for every new high school.

2.3 Descriptive Statistics

The final sample consists of more than 5 million 9th grade pupils, in about 6 000 middle schools, evenly distributed over the 8 years of observation. Among them, 57% continue in the general track, 21% go to the vocational track (all vocational tracks except for apprenticeship) and 5% repeat 9th grade. The remaining 17% are either apprentices, or dropouts. The share of pupils who follow the general track

Figure 2.2 – Tracks followed by students the year after 9th grade



Source: SCOLARITÉ data set, 9th grade pupils cohorts from 2003-2004 to 2010-2011.

is stable from 2003 to 2007 (Figure 2.2). From 2007 to 2010 we observe a slight increase. Symmetrically, the share of pupils repeating 9th grade slightly decreases at the end of the observation period. From the 2007 school year, the share of pupils going to a three-year vocational program progressively increases, while the share of pupils going to a two-year vocational program decreases.⁴ But the share of pupils who follow a vocational track remains stable over the 8 years of observation.

83 new high schools opened in mainland France over the period (Table 2.1). They represent about 2% of almost 4 000 high schools. 48 were public schools and 35 were private schools. 32 were general high schools, 27 were vocational high schools and 24 were high schools providing both vocational and general tracks. On average, around 10 new high schools opened every year over the period. Figure 2.3 shows the locations of these new high schools. They are located in municipalities with about 80 000 inhabitants on average, compared to municipalities with an average of 165 000 inhabitants for pre-existing high schools. According to Table 2.2, 10 new high schools

⁴This is mainly due to a reform of professional tracks taking place from the 2007 school year, which main objective was to encourage pupils to follow a long vocational track.

Table 2.1 – High schools openings in the sample by year and type

	Number of opening high schools					
	Total	Public	Private	LGT	LPO	LPR
2003	6	2	4	5	1	0
2004	13	9	4	5	4	4
2005	14	7	7	4	5	5
2006	12	5	7	4	3	5
2007	12	8	4	4	5	3
2008	8	7	1	3	3	2
2009	8	6	2	2	5	1
2010	10	4	6	5	1	4
Total	83	48	35	32	27	24
Mean over the period	10	6	4	4	3	3

Source: SCOLARITÉ data set, 9th grade pupils cohorts from 2003-2004 to 2010-2011.

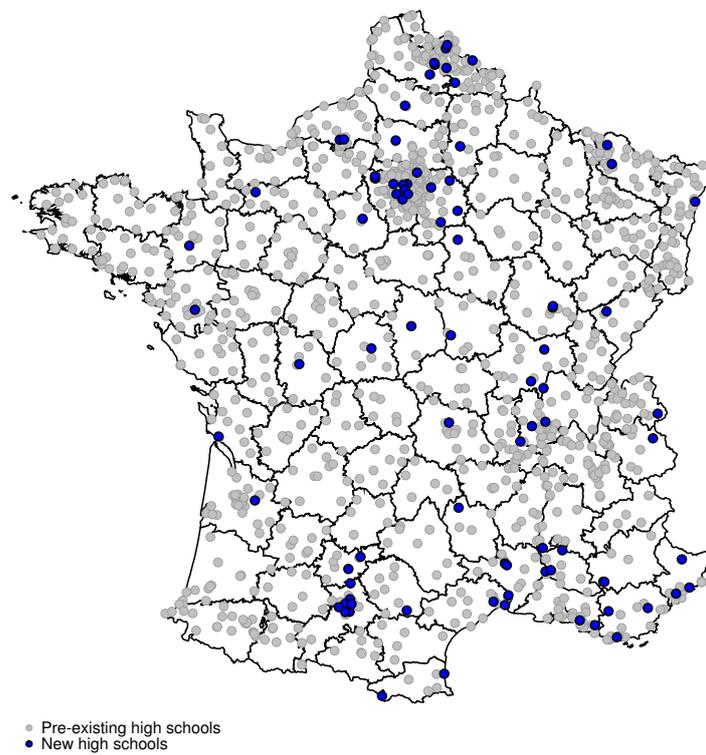
Note: LGT stands for general high schools, LP for vocational high schools and LPO are high schools that provide both vocational and general tracks.

are located in rural municipalities, they represent 12% of new high schools, compared to 5% of pre-existing high schools. 30 new high schools opened in large cities with more than 200 000 but less than 2 million inhabitants, it represents 36% of opening high schools, compared to 27% of pre-existing high schools. Thus, with respect to pre-existing high schools, new high schools seem to open more often in very small or very big municipalities. To control for this, we will use the panel nature of the data. As explained later, because schools are observed at many points in time, we do not need high schools to appear randomly on the territory.

On average, between 11 and 166 middle schools are treated each year, depending on the definition of treatment (Table 2.3). The first column of Table 2.3 shows that there are a few more closest treated middle schools than opening high schools, because some middle schools may be located at the exact same distance to a new high school.

About 55 000 pupils are enrolled in a new high school, that is, about 1.3% of pupils. Within treated middle schools, the share of pupils enrolling in a new high school the year when it opens varies between 7% on average, if we consider the median radius treated schools, and 23% on average, if we consider the closest treated

Figure 2.3 – High school openings in mainland France (2003-2010)



Source: SCOLARITÉ data set, 9th grade pupils cohorts from 2003-2004 to 2010-2011.

Table 2.2 – Types of municipalities where high schools are located

New high school	0		1	
	Freq	%	Freq	%
Municipality size				
Rural	205	5.3	10	12.0
< 5 000	160	4.2	2	2.4
< 10 000	284	7.4	3	3.6
< 20 000	372	9.7	9	10.8
< 50 000	452	11.7	10	12.0
< 100 000	441	11.5	8	9.6
< 200 000	318	8.3	4	4.8
< 2 million	1,031	26.8	30	36.1
Paris	586	15.2	7	8.4

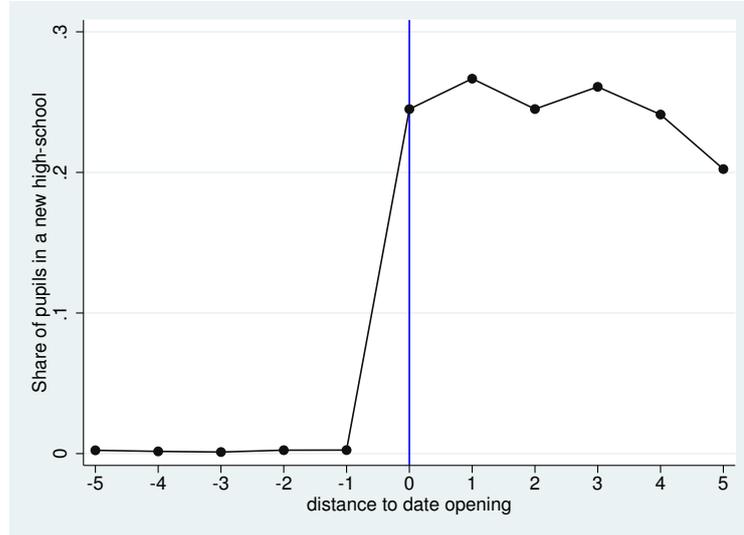
Source: SCOLARITÉ data set, 9th grade pupils cohorts from 2003-2004 to 2010-2011.

Table 2.3 – Treated middle schools in the sample by year and definition of treatment

	Number of treated middle schools			
	1st closest	2 closest	5 closest	Median
2003	6	12	30	61
2004	14	27	66	235
2005	15	29	71	165
2006	12	24	61	197
2007	13	25	62	165
2008	8	16	40	141
2009	8	17	41	188
2010	10	20	50	172
Total	86	170	421	1,324
Mean over the period	11	21	53	166

Source: SCOLARITÉ data set, 9th grade pupils cohorts from 2003-2004 to 2010-2011.

Figure 2.4 – Share of pupils entering a new high school



Source: SCOLARITÉ data set, 9th grade pupils cohorts from 2003-2004 to 2010-2011.

schools (Figure 2.4). On average every year, the 9th grade pupils of a given middle school will end up in 12 different high schools. We would then expect a new high school to attract on average 8% of a cohort.

2.4 Estimation strategy

We consider a model of repeated cross sections in which successive cohorts of 9th grade pupils are observed every year from 2003 to 2010 in S middle schools. Let \bar{Y}_{st} be the proportion of pupils enrolled in middle school s in year t who follow the track Y in year $t + 1$. We consider the following equation:

$$\bar{Y}_{st} = \alpha + \beta T_s \times \mathbf{1}\{t \geq t_s\} + \gamma' \bar{X}_{st} + \sum_{t=2003}^{2009} \delta_t \mathbf{1}_t + \sum_{s=1}^{S-1} \mu_s \mathbf{1}_s + \varepsilon_{st} \quad (2.1)$$

T_s is the treatment variable with value 1 if middle school s is treated and 0 otherwise. $\mathbf{1}\{t \geq t_s\}$ equals 1 for the years following the first year a new high school opened in

the neighbourhood of middle school s and 0 otherwise⁵. \bar{X}_{st} is a vector of pupils' characteristics in middle school s . The model includes year fixed effects, $\mathbb{1}_t$, that account for the evolution in time of track choices in the period 2003 to 2010. The middle school fixed effects control for the heterogeneity in ability and preferences across schools. The parameter of interest is β . It measures the effect of opening a new high school in the neighborhood on pupils' chosen track at the end of 9th grade.

To be able to disentangle short and long-term effects, we will then estimate the following equation:

$$\bar{Y}_{st} = \alpha + \beta_1 T_s \times \mathbb{1}\{t \geq t_s\} + \beta_2 T_s \times \mathbb{1}\{t > t_s\} + \gamma' \bar{X}_{st} + \sum_{t=2003}^{2009} \delta_t \mathbb{1}_t + \sum_{s=1}^S \mu_s \mathbb{1}_s + \varepsilon_{st} \quad (2.2)$$

where β_1 measures the instantaneous effect of a high school opening in year t_s and β_2 captures a potential additional long term effect after t_s . If β_2 is significantly different from zero, then it means that the longer term effect is different from the short term effect. Furthermore, if β_2 has the same sign as β_1 , then the effect of the treatment gets bigger over the years. If their signs are different, then the outcome tends to go back to its initial value as time goes by. Note that the parameters β do not depend on t , meaning that we suppose the effect of the treatment to be the same whatever the date when it intervenes.

The parameters β are estimated by ordinary least squares in equations (2.1) and (2.2). This estimator is equivalent to the generalized difference in differences estimator (Bertrand, Duflo, and Mullainathan, 2004; Hansen, 2007). It uses both the time and school dimensions and so accounts for potential selection into the treatment and for time trend. The middle school fixed effects control for the possibility that treated schools have unobserved characteristics correlated with high school openings. This means that high school openings need not to be exogenous events. The year fixed effects control, for instance, for the decrease in the share of pupils repeating 9th grade over the period of observation.

The difference in differences estimator relies on the assumption of common trend between the treated and the control groups. This assumption means that, if no high school opening had occurred, pupils' track choices would have evolved in the same way in treated middle schools and in non treated ones. This hypothesis cannot

⁵For the treatment definition based on the median of distance, some middle schools (around 10%) are treated twice over the observational period. In that case, we excluded observations from the year of the second opening, i.e. for these schools, $\mathbb{1}\{t \geq t_s\}$ equals 0 for the years before the first opening, 1 after the first opening, and missing starting from the year of the second opening.

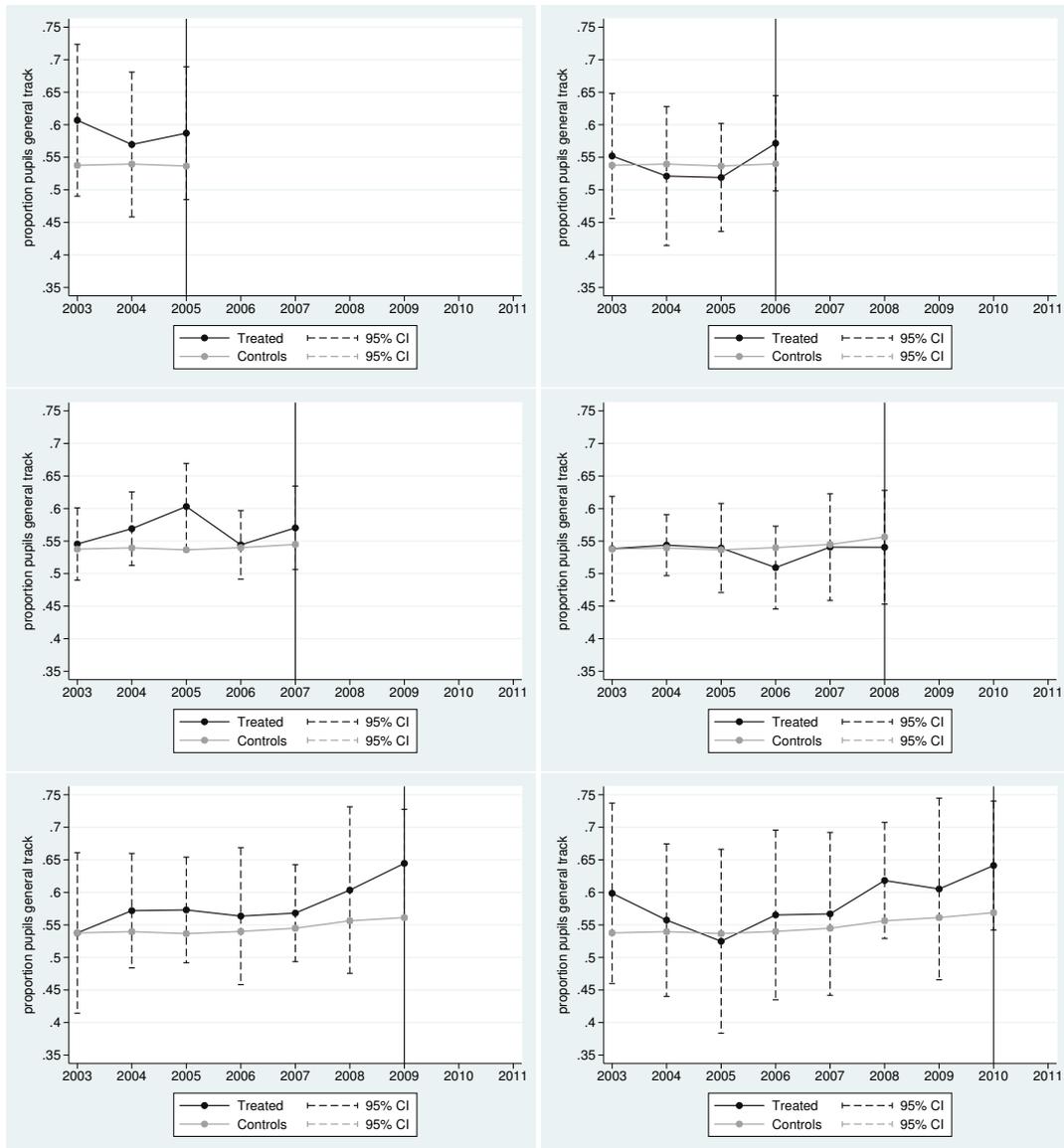
be tested directly, but the observation of the evolution in track choices in both treated and control schools before the treatment is informative. Indeed, if pupils' track choices in both groups followed a common trend before the treatment, then assuming they would have continue to evolve in similar ways if the treatment had not occurred is a credible assumption.

In our case, the period before (or after) treatment is not the same for every middle school, since new high schools may open each year. Thus, we cannot compare the treated and control groups before treatment. But we can divide the treated group (i.e schools such that $T_s = 1$) into sub-groups, by year. A given year t , treated middle schools are those with a new high school opening in their neighbourhood. Middle schools of the control group are those which are never treated over the observation period (i.e schools such that $T_s = 0$). Note that the we use the more conservative definition of the treatment (only the closest middle school to a newly opened high school is treated). For every possible year t of treatment, Figure 2.5 presents the evolution in the proportion of pupils allocated to the general track at the end of 9th grade until that date, both in the control and treatment groups. Overall, we do observe parallel trends, although the variance is of course much higher for the treated group than for the control group.

Another source of bias may be due to changes in the composition of the neighbourhood just before the treatment, that are due to the treatment. First, it could be that some parents anticipated the opening of a new high school and had their child change middle school just before the opening. If such children have unobserved characteristics correlated to preferences over tracks, then we would observe a discontinuity in allocations just before the treatment (Ashenfelter *dip*) and the common trend assumption would not hold. Second, regions may anticipate a change in pupils' preferences and decide to open a new high school to satisfy the new preferences. As we have seen, the process of opening a new high school is a long one, so that the two situations discussed here are very unlikely. As a test, we can compare the composition of schools just before and after the date of opening. Figure 2.6 presents the evolution of treated schools 9th grade cohorts composition, before and after the treatment, with respect to observable characteristics in the data, namely the proportion of boys, the proportion of each parental occupation, and the proportion of pupils benefiting from a scholarship. There is no significant discontinuity in the composition of 9th grade cohorts around the date of treatment.⁶

⁶For a more formal test of a change in treated middle schools' composition the year of the

Figure 2.5 – Evolution of the proportion of pupils allocated to the general track

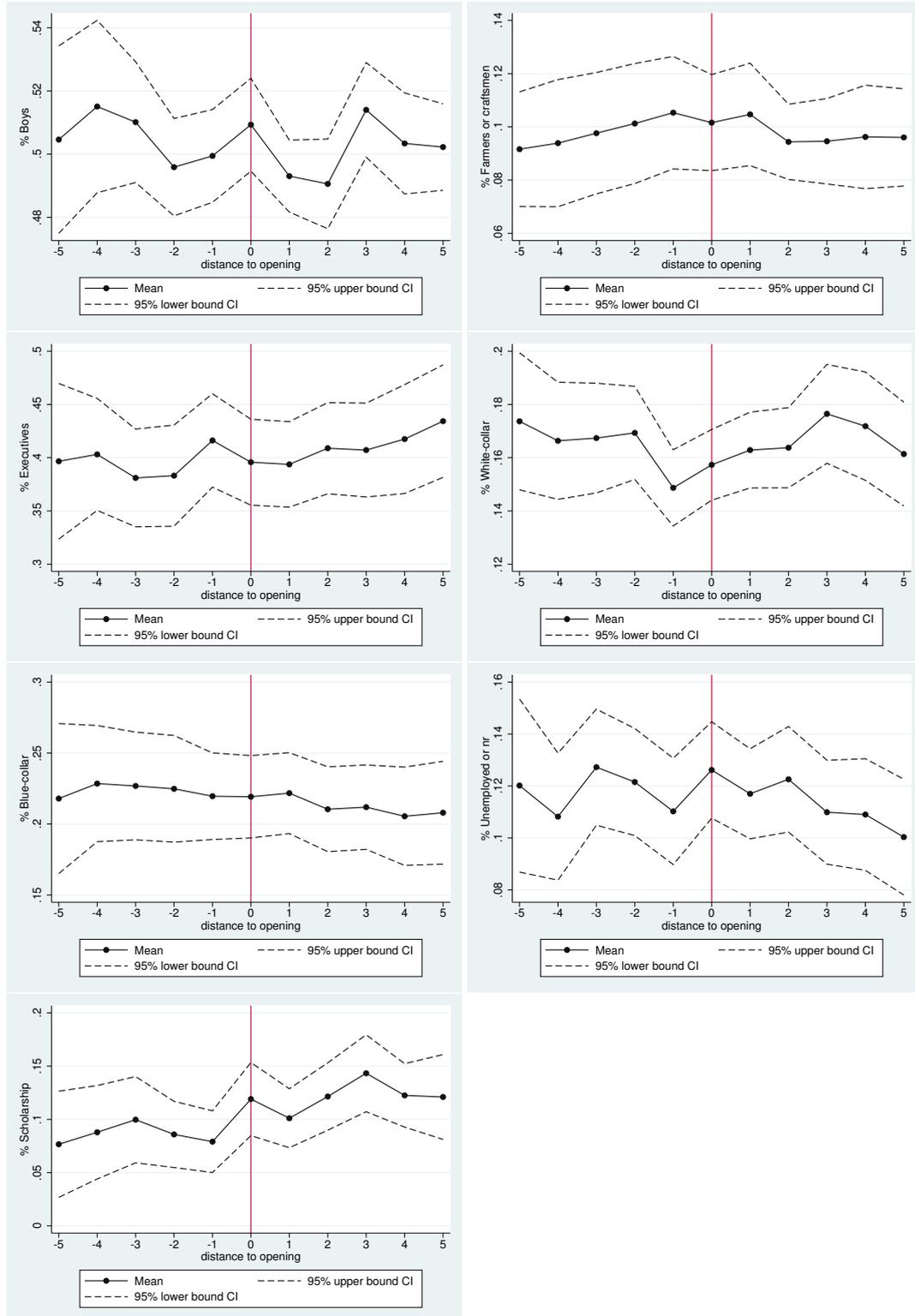


Source: SCOLARITÉ data set, 9th grade pupils cohorts from 2003-2004 to 2010-2011.

Note: Dashed lines represent 95% confidence intervals.

Lecture: Each graph plots the proportion of pupils going to the general track in treated middle schools (in black), and in non treated middle schools (in gray).

Figure 2.6 – Evolution of the composition of treated schools before and after an opening



Source: SCOLARITÉ data set, 9th grade pupils cohorts from 2003-2004 to 2010-2011.

Lecture: The first graph gives, on the y-axis, the proportion of boys every year in treated middle schools, with respect to the distance to treatment on the x-axis.

2.5 Results

Table 2.4 presents the estimates of the generalized difference in differences model presented in equation (2.1) on five outcomes: going to high school, either by attending a general track, or by attending a vocational track; repeating 9th grade; and dropping out. $T(t \geq t_s)$ represents the treatment dummy and measures the average effect of opening a new high school in treated middle schools neighborhoods. The treatment is differentiated according to the type of high school. For the moment, we present the results only for the closest treated middle school. Table 2.5 gives the results of estimating equation (2.2).

Note that we only consider the opening of public high schools here. Because the opening of a private school may be a very specific process, identifying assumptions are less likely to hold in that case. Moreover, we do not find any significant effect of opening a private high school on pupils allocation in treated schools. Table 2.8 in the appendix shows that the effects we are going to describe are completely driven by the opening of public high schools.

The opening of a new general high school (LGT) has no significant effect on the allocation of pupils of the closest middle school. $T(t \geq t_s)$ LGT is associated with a 1 to 1.7 percentage point increase in the share of pupils who access an upper secondary institution after 9th grade, but these coefficients are not significant. Opening a high school providing both general and vocational tracks (LPO) significantly increases the proportion of pupils continuing in high school from 80% to 81.6%. This is due to an increase in the proportion of pupils going to a vocational track, and a decrease in the proportion of pupils dropping out. These effects hold when covariates are added as controls (namely the proportion of boys in the middle school, the proportion of children of executives, blue-collars, etc., the proportion of pupils benefiting from a scholarship, and the school mean scores in French and in mathematics at the *Brevet* exam). In that case, the proportion of pupils attending a vocational track goes from 20.3% to 22.3% on average when a new LPO opens in the vicinity of a school. Opening a vocational high school (LPR) significantly increases the proportion of pupils continuing in a vocational track in the closest middle school, by 3.4 percentage points. This increase is simultaneous with a decrease in the proportion of pupils going to a general track, and with an increase in the proportion of repeaters, so that there is no significant impact on the proportion of pupils continuing in high school.

treatment, see Section 2.6.

In Table 2.5, we consider the dynamics of the effects. Note first that the significance tests are much less powerful, as expected, so that there is almost no significant effect. We are therefore going to focus on the effects found to be significant in Table 2.4 and look at the dynamics behind those effects. The short-run effect of opening a LPO on the proportion of pupils in vocational track is still positive but not significant. The average long-run effect has the same sign and is not significant either. The negative effect on the proportion of drop-outs, which was only marginally significant, is still negative in the short- and in the long-run, although not significant. The short-run positive effect of opening a professional high school on the proportion of repeaters is positive and significant, whereas the average long-run effect is negative. Similarly, the short-run effect on the proportion of pupils attending a vocational track is positive, but no longer significant, whereas the long-run effect is negative. If anything, the coefficients of short- and long-term effects are mainly of opposite signs. Although most coefficients are not significant, this may suggest that the instantaneous effects of opening a new high school later fade away.

To consolidate these findings, we test for the validity of the results with respect to the definition of treated middle schools. Tables 2.9 and 2.10 in the appendix give the results when the two - respectively five - closest middle schools are considered as treated. Reassuringly, the effects are qualitatively exactly similar. Only the effects are less significant. As expected, the less conservative the definition of treatment, the smaller the average treatment effects. A very robust effect is the increase in the proportion of pupils going to a vocational track when a new LPO opens. To go further, Table 2.11 presents the results when the median distance from middle schools to high schools is used to define treatment. The results are even smaller, and, when significant, they are again very similar to the previous ones. Only the opening of a new professional high school (LPR) now significantly increases the proportion of drop outs.

To sum up, our results first show that opening a new high school increases the proportion of pupils who continue in upper secondary education. This suggests that individuals are constrained by local school supply. A second and very robust result is that the effect is driven by pupils who follow a vocational track, meaning that the pupils who are constrained are those who would follow a vocational track, but repeat or drop out instead. Alleviating a supply constraint by opening a new high school allows these pupils to continue in high school.

The affected pupils thus seem to be those pupils who do not perform well enough

to access the general track but may continue in the vocational track if offered a place that matches their preferences. The observed effects may then be explained by three different mechanisms. First, opening a new high school may result in a quantitative increase in the number of vocational track positions offered. The number of pupils offered a position in vocational track will then increase and as some of them will accept the position, the share of students who continue in secondary education increases too. Second, opening a new high school may introduce new specialties of vocational studies in the neighbourhood of the pupils. Some pupils who may not have been interested in vocational studies before may then find interest in the curriculum. Third, the opening of a new high school that offers vocational tracks is likely to significantly reduce the distance between the pupils and the institution offering them a vocational track. This may result in more pupils accepting to join a vocational program. We plan on investigating those different scenarios by looking at the impact of opening a new high school on both the distance to the closest high school offering a vocational track and the theoretical number of vocational places available in high schools.

If the local school supply matters for schooling decisions such as track choices, then opening a new high school is likely to have differentiated effects depending on the *ex-ante* school supply size.

To see whether the effect differs with respect to the size of local high school supply, we constructed a variable approximating each middle school's "potential" supply. More precisely, for every middle school, and every year, we computed the number of high schools (excluding new high schools), weighted by the distance to the middle school. This was calculated separately for each type of high school, so that every middle school has a measure of private high school potential supply, general high school potential supply, vocational high school potential supply, etc. Then the treatment variable was interacted with a dummy which equals one if the treated school lies within the top half of the distribution of potential supply, and zero otherwise. The idea is to analyze whether the impact of opening a new high school is higher in neighborhoods where the supply is relatively low (i.e. below the median).

Table 2.6 gives the results of the effect of opening a new high school interacted with the private high school potential supply measure. The opening of a new LPO high school is not affected by the size of local private supply; the increase in the proportion of pupils going to a vocational track in treated schools is robust, and is

the same whatever the size of the private potential supply. The effect of opening a new professional school (LPR), on the other hand, differs with respect to the number of potential private high schools. Opening a new LPR high school still increases the share of pupils continuing in vocational high school, in neighborhoods where the potential private supply is below the median. However, when the treated school has a large potential private supply, then the effect of opening a new professional high school almost vanishes. The proportion of repeaters increases with the opening of a new professional high school, and the increase is bigger when the local private sector supply is larger. This may suggest that, when the potential private supply is large, pupils repeat and try again to continue in high school the following year.

Because the *Ile de France* region is characterized by a larger and more diverse school supply than other regions, we estimated separate effects for this particular region. Over 83 new high schools, 12 opened in *Ile de France* over the period, among them, 8 were public high schools. Note that only 1 private professional high school opened in *Ile de France* over the period. Table 2.7 presents the results distinguishing openings that occurred in *Ile de France* ($idf=1$) from other high school openings. The effects are robust to the exclusion of *Ile de France*. We do not find any significant effect of opening a new public high school in *Ile de France*, although the number of openings is of course very small. Only opening a new LGT in *Ile de France* significantly decreases the proportion of pupils going to a vocational track.

Table 2.4 – DID estimates of the effect of opening a new high school on track choice

	%All tracks		%High school			%Repetition		%Dropout		
			%General		%Vocational					
T(t >= ts) LGT	0.017 (0.016)	0.010 (0.015)	0.018 (0.012)	0.004 (0.010)	-0.001 (0.011)	0.006 (0.011)	-0.006 (0.004)	-0.005 (0.004)	-0.011 (0.016)	-0.005 (0.016)
T(t >= ts) LPO	0.016* (0.010)	0.016* (0.009)	-0.003 (0.010)	-0.004 (0.011)	0.019*** (0.006)	0.020*** (0.007)	0.003 (0.006)	0.003 (0.006)	-0.019* (0.010)	-0.019* (0.010)
T(t >= ts) LPR	-0.012 (0.018)	-0.015 (0.018)	-0.043* (0.025)	-0.049** (0.021)	0.032* (0.017)	0.034*** (0.013)	0.031*** (0.004)	0.031*** (0.005)	-0.019 (0.018)	-0.016 (0.018)
Intercept	0.800*** (0.001)	0.809*** (0.008)	0.593*** (0.001)	0.605*** (0.009)	0.208*** (0.001)	0.203*** (0.007)	0.040*** (0.000)	0.057*** (0.004)	0.160*** (0.001)	0.134*** (0.007)
Controls	No	Yes								
Year fe.	Yes									
R2	0.63	0.64	0.82	0.84	0.81	0.82	0.45	0.46	0.65	0.66
Nbr obs	53,182	53,045	53,182	53,045	53,182	53,045	53,182	53,045	53,182	53,045
Nbr clusters	6,803	6,785	6,803	6,785	6,803	6,785	6,803	6,785	6,803	6,785

85

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Table 2.5 – DID estimates of the short- and long-term effect of opening a new high school

	%High school						%Repetition	%Dropout		
	%All tracks		%General		%Vocational					
T(t >= ts) LGT	0.022 (0.015)	0.018 (0.015)	0.018 (0.015)	0.011 (0.011)	0.004 (0.013)	0.007 (0.011)	-0.011* (0.006)	-0.010 (0.006)	-0.011 (0.015)	-0.008 (0.015)
T(t > ts) LGT	0.013 (0.010)	0.008 (0.010)	0.007 (0.014)	-0.005 (0.011)	0.006 (0.014)	0.013 (0.012)	0.004 (0.007)	0.005 (0.006)	-0.017 (0.013)	-0.013 (0.014)
T(t >= ts) LPO	0.014 (0.011)	0.013 (0.010)	0.005 (0.013)	0.004 (0.012)	0.008 (0.011)	0.010 (0.011)	-0.002 (0.009)	-0.002 (0.009)	-0.012 (0.010)	-0.012 (0.009)
T(t > ts) LPO	-0.004 (0.013)	-0.006 (0.011)	-0.021 (0.017)	-0.025 (0.016)	0.017 (0.013)	0.019 (0.013)	0.008 (0.010)	0.008 (0.010)	-0.004 (0.008)	-0.002 (0.007)
T(t >= ts) LPR	-0.053* (0.032)	-0.046 (0.028)	-0.094 (0.068)	-0.079 (0.058)	0.041 (0.045)	0.033 (0.039)	0.046*** (0.014)	0.043*** (0.013)	0.007 (0.030)	0.003 (0.029)
T(t > ts) LPR	0.021 (0.032)	0.015 (0.024)	0.046 (0.050)	0.032 (0.040)	-0.025 (0.024)	-0.017 (0.022)	-0.014 (0.022)	-0.013 (0.021)	-0.007 (0.025)	-0.002 (0.024)
Intercept	0.800*** (0.001)	0.824*** (0.008)	0.593*** (0.001)	0.616*** (0.009)	0.207*** (0.001)	0.209*** (0.007)	0.040*** (0.000)	0.055*** (0.004)	0.160*** (0.001)	0.121*** (0.007)
Controls	No	Yes								
Year fe.	Yes									
R2	0.64	0.65	0.82	0.85	0.82	0.83	0.47	0.47	0.66	0.67
Nbr obs	46,375	46,279	46,375	46,279	46,375	46,279	46,375	46,279	46,375	46,279
Nbr clusters	6,763	6,746	6,763	6,746	6,763	6,746	6,763	6,746	6,763	6,746

98

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Table 2.6 – DID estimates of the effect of opening a new high school depending on local private supply

	%High school						%Repetition		%Dropout	
	%All tracks		%General		%Vocational					
T(t >= ts) LGT	0.005 (0.017)	-0.000 (0.015)	-0.002 (0.018)	-0.013 (0.014)	0.008 (0.013)	0.013 (0.014)	-0.004 (0.008)	-0.003 (0.008)	-0.001 (0.016)	0.003 (0.015)
T(t >= ts) LPO	0.017 (0.011)	0.017* (0.010)	-0.002 (0.009)	-0.003 (0.011)	0.020** (0.008)	0.020** (0.010)	0.008 (0.007)	0.008 (0.007)	-0.025* (0.014)	-0.025* (0.014)
T(t >= ts) LPR	0.005 (0.020)	0.004 (0.016)	-0.041 (0.036)	-0.041 (0.030)	0.046*** (0.016)	0.045*** (0.014)	0.027*** (0.004)	0.027*** (0.005)	-0.032 (0.023)	-0.031 (0.019)
T(t >= ts) LGT × MP Pri high	0.020 (0.029)	0.017 (0.027)	0.035 (0.023)	0.029* (0.018)	-0.015 (0.022)	-0.012 (0.022)	-0.004 (0.009)	-0.003 (0.009)	-0.016 (0.030)	-0.014 (0.028)
T(t >= ts) LPO × MP Pri high	-0.003 (0.021)	-0.003 (0.020)	-0.002 (0.023)	-0.002 (0.024)	-0.001 (0.010)	-0.002 (0.012)	-0.013 (0.012)	-0.013 (0.012)	0.016 (0.018)	0.017 (0.017)
T(t >= ts) LPR × MP Pri high	-0.052*** (0.020)	-0.059*** (0.016)	-0.007 (0.036)	-0.026 (0.030)	-0.046*** (0.016)	-0.033** (0.014)	0.012*** (0.004)	0.011** (0.005)	0.041* (0.023)	0.047** (0.019)
Intercept	0.800*** (0.001)	0.809*** (0.008)	0.593*** (0.001)	0.606*** (0.009)	0.208*** (0.001)	0.203*** (0.007)	0.040*** (0.000)	0.057*** (0.004)	0.160*** (0.001)	0.134*** (0.007)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year fe.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.63	0.64	0.82	0.84	0.81	0.82	0.45	0.46	0.65	0.66
Nbr obs	53,182	53,045	53,182	53,045	53,182	53,045	53,182	53,045	53,182	53,045
Nbr clusters	6,803	6,785	6,803	6,785	6,803	6,785	6,803	6,785	6,803	6,785

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Table 2.7 – Separate estimates of the effect of opening a new high school for Ile de France

	%All tracks		%High school			%Repetition		%Dropout		
			%General		%Vocational					
T(t >= ts) LGT	0.019 (0.020)	0.013 (0.018)	0.012 (0.015)	-0.001 (0.012)	0.007 (0.013)	0.014 (0.013)	-0.005 (0.005)	-0.003 (0.005)	-0.015 (0.021)	-0.009 (0.020)
T(t >= ts) LPO	0.019* (0.010)	0.020** (0.009)	-0.001 (0.011)	-0.000 (0.011)	0.021*** (0.007)	0.020*** (0.008)	0.001 (0.007)	0.001 (0.007)	-0.020* (0.012)	-0.021* (0.011)
T(t >= ts) LPR	-0.012 (0.018)	-0.015 (0.018)	-0.043* (0.025)	-0.049** (0.021)	0.032* (0.017)	0.034*** (0.013)	0.031*** (0.004)	0.031*** (0.005)	-0.019 (0.018)	-0.016 (0.018)
T(t >= ts) LGT=1 × idf=1	-0.010 (0.023)	-0.013 (0.022)	0.029 (0.022)	0.024 (0.017)	-0.039** (0.017)	-0.037** (0.015)	-0.007 (0.007)	-0.007 (0.008)	0.017 (0.022)	0.020 (0.021)
T(t >= ts) LPO=1 × idf=1	-0.021 (0.018)	-0.027 (0.018)	-0.011 (0.018)	-0.023 (0.026)	-0.011 (0.007)	-0.004 (0.013)	0.012 (0.012)	0.012 (0.011)	0.010 (0.014)	0.015 (0.016)
T(t >= ts) LPR=1 × idf=1	0.000 (.)									
Intercept	0.800*** (0.001)	0.809*** (0.008)	0.593*** (0.001)	0.606*** (0.009)	0.208*** (0.001)	0.203*** (0.007)	0.040*** (0.000)	0.057*** (0.004)	0.160*** (0.001)	0.134*** (0.007)
Controls	No	Yes								
Year fe.	Yes									
R2	0.63	0.64	0.82	0.84	0.81	0.82	0.45	0.46	0.65	0.66
Nbr obs	53,182	53,045	53,182	53,045	53,182	53,045	53,182	53,045	53,182	53,045
Nbr clusters	6,803	6,785	6,803	6,785	6,803	6,785	6,803	6,785	6,803	6,785

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

2.6 Robustness

The common trend assumption requires that, in the absence of treatment, treated schools would not have evolved differently from control schools. Although this hypothesis is impossible to test directly, we can check the robustness of our results to some changes in the specification of the model.

First, we excluded schools which are never treated over the period. The reason for this robustness test is that the 86 treated middle schools may be a very specific population; assuming that their evolution in the absence of treatment would be comparable to that of the rest of the entire population of French middle schools may be a strong assumption. One possibility is to restrict the sample to middle schools being treated at some point over the observation period. Such schools are likely to be very similar, the only difference being the date when they are treated. The drawbacks are that one observational period is lost (i.e. middle schools being treated the last year serve as "control" group ($T_s = 0$)), and that the sample size is much smaller. The results are given in Table 2.12. Although the number of observations shrinks from over 53 000 to less than 700, the results are very similar to that of Table 2.4.

Second, we added a linear trend in the regression. Suppose that would-be treated schools follow a specific trend in the outcomes before the treatment. Then it may be that we wrongly attribute the effects to a change in high school supply, while they are actually driven by a trend. If this is the case, then the effects should vanish, or change, once a linear trend is added in the regression. To test for this, we defined the distance to the date of opening in the sub-sample of schools being treated over the period, and we added this variable to the regression. Table 2.13 in the appendix gives the results. For each outcome, the results without and with the trend are given, to make the comparison easier. Whether or not we control for a trend, the results are almost identical.

Third, we tested for a change in the social composition of treated schools at the exact date of the opening of a new high school. As explained earlier, we need the school composition to have not change just before the treatment, so that we would not be able to separate the treatment effect from a modification of the treated population. To formally test this, we regressed equation (2.2) on the observable social characteristics at the school level: the proportion of boys; the proportion of children of farmers and craftsmen, executives, white-collars, blue-collars, or unemployed; the proportion of pupils benefiting from a scholarship. Table 2.14 in the appendix gives

the results and shows that $T(t \geq t_s)$ is never significant, meaning that there is no discontinuity in the social composition of treated schools the year of the treatment.

Lastly, we ran a Placebo test, assuming that the treatment took place two years before the actual date of treatment. If the common trend assumption holds, there should not be any significant effect of the treatment two years before it actually happens. To test for this, we thus ran the exact same difference in differences regression, assuming that treated middle schools were treated two years earlier, and controlling for the trend after the actual date of treatment. Table 2.15 in the appendix gives the results. If we move the treatment two years back in time, there is no significant effect of the opening of a new LGT high school nor of a new LPO high school. The opening of a new professional high school has a marginally significant effect on the proportion of pupils continuing in high school. After the actual date of treatment, the treatment effects are similar to that of Table 2.4.

2.7 Conclusion

This paper aims at analysing the causal effect of a change in local school supply on pupils track choice at the end of lower secondary education. We take advantage of high school openings to highlight the constraint school supply exerts on individual schooling choices. We use an exceptionally rich data set in which we observe every pupil enrolled in 9th grade in mainland France every year from 2002 to 2011. From the data, we recover the information about new high schools each school year. A model of generalized difference in differences makes use of the variation in time and location of opening high schools to identify the causal effect of a change in local school supply on the allocation of pupils at the end of middle school.

We show that pupils are constrained by the local school supply since opening a new high school increases the proportion of pupils who continue in upper secondary education by up to 1.6 percentage point. This is due to pupils who go to a vocational track. The effect is driven by the opening of either high schools that offer vocational tracks (LPR) or high schools that offer both vocational and general tracks (LPO). This increase comes with a decrease in the share of dropouts. The effect is mainly instantaneous and seems to fade over time. Those results are robust to various definitions of the treated group. They also persist when we introduce control variables and add a linear trend. Selecting a smaller group of control middle schools does not change the results either.

The magnitude of the effect is small, but not easy to compare to the existing literature. First, the effect of opening a new school varies a lot across studies and countries. For instance, building a new school increases the primary education enrolment rate by 0.3 percentage point in Mozambique (Handa, 2002) and by 35 to 52 percentage points in Afghanistan (Burde and Linden, 2013). Furthermore, the expected magnitude is of course not to be the same in developing and in developed countries. Second, we don't expect to find the same magnitude in primary and in secondary education. Third, to our knowledge, there is no pre-existing study of the effect of opening a new school on enrolment in upper secondary education. Dickerson and McIntosh (2013) setting is very similar to ours, although they look at the effect of distance to education institutions on post-compulsory secondary education, and not that of the opening of a new school. Following the results of Goux, Gurgand, and Maurin (2015), our findings suggest that opening new high schools that offer vocational tracks may improve students achievement.

These preliminary results call for further investigation of the type of constraint pupils are facing. We would like to look first at the impact of the new openings on the distance to high school and more specifically the distance to high schools offering a vocational program. We anticipate a significant reduction of those distances and would like to look at the effect of these reductions on the allocation of students. We also plan on looking at the link between opening a new high school and the supply of secondary education. We will do so by estimating the effect of opening a new high school on the number of sits available in each type of track. We will then look at how the changes in the number of sits available impact the allocation of pupils. Those next steps should allow us to tell more about the relationship between local supply of schooling and schooling decisions.

Appendix

Table 2.8 – DID estimates of the effect of opening a new high school by sector

	%High school						%Repetition		%Dropout	
	%All tracks		%General		%Vocational					
T(t >= ts) public	0.014*	0.011	0.003	-0.004	0.011*	0.015**	0.001	0.002	-0.015*	-0.012
	(0.009)	(0.008)	(0.008)	(0.007)	(0.006)	(0.006)	(0.004)	(0.004)	(0.009)	(0.009)
T(t >= ts) private	-0.004	-0.002	-0.003	0.001	-0.001	-0.003	0.001	0.001	0.003	0.002
	(0.006)	(0.006)	(0.009)	(0.009)	(0.007)	(0.006)	(0.004)	(0.004)	(0.006)	(0.006)
Intercept	0.801***	0.809***	0.593***	0.605***	0.208***	0.204***	0.040***	0.057***	0.160***	0.134***
	(0.001)	(0.008)	(0.001)	(0.009)	(0.001)	(0.007)	(0.000)	(0.004)	(0.001)	(0.007)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year fe.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.63	0.64	0.82	0.84	0.81	0.82	0.45	0.46	0.65	0.66
Nbr obs	53,182	53,045	53,182	53,045	53,182	53,045	53,182	53,045	53,182	53,045
Nbr clusters	6,803	6,785	6,803	6,785	6,803	6,785	6,803	6,785	6,803	6,785

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Table 2.9 – DID estimates of the effect of opening a new high school on the two closest schools

	%High school						%Repetition	%Dropout		
	%All tracks		%General		%Vocational					
T(t >= ts) LGT	0.016 (0.011)	0.011 (0.010)	0.017* (0.009)	0.007 (0.008)	-0.001 (0.006)	0.003 (0.006)	-0.003 (0.003)	-0.002 (0.003)	-0.013 (0.011)	-0.009 (0.010)
T(t >= ts) LPO	0.016* (0.008)	0.013 (0.008)	0.004 (0.009)	-0.002 (0.009)	0.012** (0.006)	0.015** (0.006)	-0.003 (0.005)	-0.002 (0.005)	-0.012* (0.007)	-0.011 (0.007)
T(t >= ts) LPR	-0.011 (0.011)	-0.005 (0.012)	-0.022 (0.020)	-0.012 (0.022)	0.012 (0.015)	0.006 (0.015)	0.005 (0.010)	0.004 (0.010)	0.006 (0.013)	0.002 (0.011)
Intercept	0.800*** (0.001)	0.809*** (0.008)	0.593*** (0.001)	0.605*** (0.009)	0.208*** (0.001)	0.203*** (0.007)	0.040*** (0.000)	0.057*** (0.004)	0.160*** (0.001)	0.134*** (0.007)
Controls	No	Yes								
Year fe.	Yes									
R2	0.63	0.64	0.82	0.84	0.81	0.82	0.45	0.46	0.65	0.66
Nbr obs	53,175	53,038	53,175	53,038	53,175	53,038	53,175	53,038	53,175	53,038
Nbr clusters	6,803	6,785	6,803	6,785	6,803	6,785	6,803	6,785	6,803	6,785

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Table 2.10 – DID estimates of the effect of opening a new high school on the five closest schools

	%High school						%Repetition		%Dropout	
	%All tracks		%General		%Vocational					
T(t >= ts) LGT	0.003 (0.006)	-0.001 (0.006)	0.005 (0.006)	-0.002 (0.005)	-0.002 (0.004)	0.001 (0.004)	0.001 (0.002)	0.002 (0.002)	-0.004 (0.006)	-0.001 (0.005)
T(t >= ts) LPO	0.009 (0.006)	0.007 (0.005)	-0.001 (0.006)	-0.004 (0.006)	0.010** (0.004)	0.011*** (0.004)	0.003 (0.003)	0.004 (0.003)	-0.012** (0.005)	-0.011** (0.005)
T(t >= ts) LPR	-0.006 (0.009)	-0.004 (0.008)	-0.018 (0.013)	-0.015 (0.012)	0.013 (0.009)	0.011 (0.008)	-0.001 (0.007)	-0.002 (0.007)	0.007 (0.008)	0.006 (0.007)
Intercept	0.800*** (0.001)	0.809*** (0.008)	0.593*** (0.001)	0.605*** (0.009)	0.208*** (0.001)	0.203*** (0.007)	0.040*** (0.000)	0.057*** (0.004)	0.160*** (0.001)	0.134*** (0.007)
Controls	No	Yes								
Year fe.	Yes									
R2	0.63	0.64	0.82	0.84	0.81	0.82	0.45	0.46	0.65	0.66
Nbr obs	53,105	52,974	53,105	52,974	53,105	52,974	53,105	52,974	53,105	52,974
Nbr clusters	6,800	6,782	6,800	6,782	6,800	6,782	6,800	6,782	6,800	6,782

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Table 2.11 – DID estimates of the effect of opening a new high school within median radius

	%All tracks		%High school			%Repetition		%Dropout		
			%General		%Vocational					
T(t >= ts) LGT	0.009 (0.006)	0.004 (0.006)	0.012** (0.006)	0.003 (0.005)	-0.003 (0.005)	0.002 (0.005)	0.002 (0.002)	0.003 (0.002)	-0.011* (0.006)	-0.007 (0.005)
T(t >= ts) LPO	0.009*** (0.003)	0.006** (0.003)	0.005 (0.004)	-0.001 (0.003)	0.004 (0.003)	0.007*** (0.003)	-0.004** (0.002)	-0.003* (0.002)	-0.005** (0.003)	-0.003 (0.002)
T(t >= ts) LPR	-0.019** (0.008)	-0.016** (0.008)	-0.013 (0.010)	-0.008 (0.008)	-0.006 (0.008)	-0.009 (0.007)	0.000 (0.005)	-0.001 (0.005)	0.018*** (0.007)	0.017** (0.007)
Intercept	0.800*** (0.001)	0.811*** (0.008)	0.592*** (0.001)	0.607*** (0.009)	0.208*** (0.001)	0.204*** (0.007)	0.040*** (0.000)	0.057*** (0.004)	0.160*** (0.001)	0.132*** (0.007)
Controls	No	Yes								
Year fe.	Yes									
R2	0.62	0.64	0.82	0.84	0.81	0.82	0.45	0.46	0.65	0.66
Nbr obs	52,157	52,033	52,157	52,033	52,157	52,033	52,157	52,033	52,157	52,033
Nbr clusters	6,776	6,759	6,776	6,759	6,776	6,759	6,776	6,759	6,776	6,759

95

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Table 2.12 – DID estimates of the effect of opening a new high school without "super" control group

	%High school						%Repetition	%Dropout		
	%All tracks		%General		%Vocational					
T(t >= ts) LGT	0.019 (0.016)	0.008 (0.014)	0.019 (0.013)	0.001 (0.010)	0.001 (0.012)	0.007 (0.012)	-0.011** (0.005)	-0.007 (0.005)	-0.009 (0.017)	-0.001 (0.015)
T(t >= ts) LPO	0.018* (0.010)	0.019* (0.010)	-0.003 (0.010)	-0.002 (0.012)	0.021*** (0.007)	0.022*** (0.008)	-0.001 (0.007)	-0.001 (0.007)	-0.018* (0.010)	-0.018* (0.010)
T(t >= ts) LPR	-0.009 (0.021)	-0.022 (0.015)	-0.045 (0.027)	-0.055** (0.021)	0.035* (0.018)	0.033** (0.014)	0.028*** (0.005)	0.029*** (0.005)	-0.018 (0.020)	-0.007 (0.015)
Intercept	0.802*** (0.007)	0.854*** (0.053)	0.625*** (0.009)	0.585*** (0.066)	0.176*** (0.007)	0.269*** (0.055)	0.047*** (0.004)	0.057* (0.032)	0.151*** (0.006)	0.089* (0.048)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Year fe.	Yes	Yes								
R2	0.71	0.74	0.85	0.88	0.86	0.87	0.44	0.46	0.71	0.73
Nbr obs	689	689	689	689	689	689	689	689	689	689
Nbr clusters	90	90	90	90	90	90	90	90	90	90

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Table 2.13 – DID estimates of the effect of opening a new high school with pre-treatment trend

	%All tracks		%High school			%Repetition		%Dropout		
			%General		%Vocational					
T(t >= ts) LGT	0.008 (0.014)	0.008 (0.014)	0.001 (0.010)	0.001 (0.010)	0.007 (0.012)	0.007 (0.012)	-0.007 (0.005)	-0.006 (0.005)	-0.001 (0.015)	-0.001 (0.015)
T(t >= ts) LPO	0.019* (0.010)	0.019* (0.010)	-0.002 (0.012)	-0.002 (0.012)	0.022*** (0.008)	0.021*** (0.008)	-0.001 (0.007)	-0.001 (0.007)	-0.018* (0.010)	-0.018* (0.010)
T(t >= ts) LPR	-0.022 (0.015)	-0.023 (0.015)	-0.055** (0.021)	-0.055** (0.021)	0.033** (0.014)	0.032** (0.014)	0.029*** (0.005)	0.030*** (0.005)	-0.007 (0.015)	-0.007 (0.015)
Intercept	0.854*** (0.053)	0.846*** (0.055)	0.585*** (0.066)	0.568*** (0.067)	0.269*** (0.055)	0.278*** (0.057)	0.057* (0.032)	0.058* (0.032)	0.089* (0.048)	0.096* (0.050)
Controls	Yes	Yes	Yes							
Year fe.	Yes	Yes	Yes							
Trend	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
R2	0.74	0.73	0.88	0.87	0.87	0.87	0.46	0.45	0.73	0.73
Nbr obs	689	672	689	672	689	672	689	672	689	672
Nbr clusters	90	87	90	87	90	87	90	87	90	87

97

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Table 2.14 – DID estimates of the effect of opening a new high school on school social composition

	% Boys	% Farmers, craftsmen	% Executives	% White-collar	% Blue-collar	% Unemployed or nr	% Scholarship
T(t >= ts) LGT	0.002 (0.012)	0.002 (0.008)	-0.015 (0.012)	0.010 (0.013)	0.001 (0.013)	0.002 (0.007)	0.017 (0.022)
T(t > ts) LGT	0.006 (0.011)	0.010 (0.008)	0.016 (0.017)	-0.010 (0.013)	-0.017* (0.010)	0.001 (0.007)	-0.044*** (0.012)
T(t >= ts) LPO	-0.013 (0.014)	0.009 (0.007)	-0.007 (0.013)	0.004 (0.011)	-0.008 (0.010)	0.003 (0.012)	-0.051 (0.037)
T(t > ts) LPO	-0.001 (0.013)	0.003 (0.007)	0.013 (0.012)	-0.007 (0.011)	-0.002 (0.008)	-0.008 (0.011)	0.011 (0.031)
T(t >= ts) LPR	-0.041 (0.033)	0.001 (0.025)	0.023 (0.044)	-0.015 (0.019)	-0.007 (0.025)	-0.002 (0.022)	0.058 (0.085)
T(t > ts) LPR	0.020 (0.017)	-0.004 (0.014)	0.029 (0.035)	0.002 (0.014)	-0.020 (0.022)	-0.007 (0.024)	0.014 (0.053)
Intercept	0.497*** (0.001)	0.105*** (0.000)	0.370*** (0.001)	0.162*** (0.001)	0.263*** (0.001)	0.101*** (0.001)	0.142*** (0.001)
Year fe.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.44	0.82	0.93	0.71	0.90	0.83	0.59
Nbr obs	46,375	46,356	46,356	46,356	46,356	46,356	46,356
Nbr clusters	6,763	6,760	6,760	6,760	6,760	6,760	6,760

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Table 2.15 – DID estimates of the Placebo effect of opening a new high school two years before

	%High school						%Repetition		%Dropout	
	%All tracks		%General		%Vocational					
T(t >= ts-2) LGT	0.002 (0.011)	0.001 (0.012)	0.003 (0.013)	-0.001 (0.014)	-0.001 (0.011)	0.001 (0.011)	-0.005 (0.007)	-0.006 (0.007)	0.003 (0.013)	0.005 (0.014)
T(t >= ts-2) LPO	-0.014 (0.021)	-0.013 (0.020)	-0.022 (0.021)	-0.019 (0.019)	0.008 (0.007)	0.005 (0.008)	0.008 (0.008)	0.008 (0.008)	0.006 (0.017)	0.005 (0.015)
T(t >= ts-2) LPR	0.055* (0.031)	0.053* (0.031)	0.037 (0.034)	0.039 (0.036)	0.018 (0.034)	0.014 (0.035)	-0.008 (0.010)	-0.009 (0.010)	-0.047 (0.035)	-0.044 (0.034)
T(t >= ts) LGT	0.002 (0.016)	-0.005 (0.016)	0.010 (0.014)	-0.004 (0.012)	-0.008 (0.013)	-0.001 (0.013)	0.003 (0.005)	0.005 (0.004)	-0.005 (0.017)	0.000 (0.017)
T(t >= ts) LPO	0.019** (0.010)	0.020** (0.010)	0.002 (0.011)	0.005 (0.010)	0.017** (0.007)	0.015** (0.006)	0.001 (0.008)	0.001 (0.009)	-0.020* (0.010)	-0.021** (0.010)
T(t >= ts) LPR	-0.040 (0.033)	-0.045 (0.029)	-0.092 (0.060)	-0.103** (0.049)	0.053 (0.035)	0.058** (0.028)	0.027*** (0.006)	0.028*** (0.006)	0.012 (0.037)	0.017 (0.034)
Intercept	0.790*** (0.001)	0.795*** (0.009)	0.582*** (0.001)	0.579*** (0.010)	0.209*** (0.001)	0.216*** (0.008)	0.046*** (0.000)	0.062*** (0.005)	0.164*** (0.001)	0.143*** (0.008)
Controls	No	Yes								
Year fe.	Yes									
R2	0.65	0.67	0.83	0.86	0.82	0.83	0.48	0.49	0.68	0.69
Nbr obs	39,621	39,521	39,621	39,521	39,621	39,521	39,621	39,521	39,621	39,521
Nbr clusters	6,733	6,715	6,733	6,715	6,733	6,715	6,733	6,715	6,733	6,715

Source: SCOLARITÉ data set, cohorts of pupils in 9th grade in 2003-2010.

Note: *** p-value<0.001, ** p-value<0.05, * p-value<0.1. All estimations use year and middle school fixed effects. Observations are weighted by the number of 9th grade pupils in the school. Standard errors in parenthesis account for the autocorrelation of the residuals between observations of the same middle school.

Chapter 3

Good teaching and good grades. Can you buy pedagogy?

This chapter is based on joint work with Ronan LE SAOUT

Introduction

A growing part of post-secondary education institutions use student evaluation of teaching (SET). According to a survey by Becker, Bosshardt, and Watts (2012), 93% of American departments of economics reported using SET to evaluate teaching in 2011. In France, this practice was introduced only recently and is much more heterogenous.¹ Many universities and institutions don't have such evaluations, while some have been organizing them for several years. Within universities, the use of SET varies a lot across departments; it is more prevalent in scientific departments and in vocational tracks than in general ones. Moreover SET seems to be more frequent in selective tracks.²

The primary objective of SET is for the administration to assess the quality of teaching. On average, it accounts for 50% of American instructors' teaching evaluation, according to Becker, Bosshardt, and Watts (2012). Thus SET plays a

¹A 1997 decree stipulates that every French higher education institution delivering a university diploma should evaluate teaching, and that this evaluation should take students judgment into account.

²So far and to our knowledge, there is no quantitative evaluation of the use of SET in French higher education institutions. For a qualitative review, see for instance the report for the "Haut Conseil de l'évaluation de l'école", Dejean (2002).

significant role in teachers' hiring and promotion procedures. However, the validity of SET as a good measure of teaching quality is highly controversial. The prime question is to know whether students are able to correctly evaluate the quality of teaching. Student evaluations seem to be a reliable measure, in the sense that inter-rater reliability is high (i.e. evaluations of the same course by two different students are highly correlated). They are stable over time, and they are relatively highly correlated with other measures of teaching quality self-evaluation, peer evaluation, etc. Marsh and Roche, 1997. Moreover, student evaluations perform better than objective characteristics such as teachers' salary and status in explaining students' achievement (Hoffmann and Oreopoulos, 2009). But they may not measure all the dimensions that constitute good teaching (d'Apollonia and Abrami, 1997; Greenwald, 1997).

From a theoretical point of view, SET can be seen as a principal-agent-client framework in which teachers' effort and pedagogical qualities cannot be fully observed by the administration (Klitgaard, 1990). Teachers can get good evaluations either by actually improving teaching or by "corrupting" (i.e. giving good grades in order to get good evaluations). Braga, Paccagnella, and Pellizzari (2011) consider, for instance, a model in which teachers choose between two types of teaching activities: real teaching or teaching-to-the-test. Since the latter requires less effort from students than the former, good teachers may receive bad evaluations if teaching-to-the-test is effective. An alternative theory (Franz, 2010) is that teachers may give good grades to prevent students from pestering for better grades.

Thus, if most empirical studies find a positive correlation between grades and student evaluations, it is difficult to assess a causal link. First, good teachers should cause good grades and get higher evaluations at the same time. In that case, grades and evaluations are simultaneously determined. Second, students may select themselves into courses they like and they are good at, and consequently they may give higher evaluations (or, on the contrary, they may evaluate more severely if their learning expectations are high). This would result in a selection bias. Third, unobserved teachers' characteristics (such as charisma, clarity, enthusiasm, etc.) are undoubtedly related to students' achievement and evaluations at the same time, which may create endogeneity. To rule out unobserved teacher's characteristics effect, Ewing (2012) and Isely and Singh (2005) include teacher's fixed effects and find a positive effect of expected grades on evaluations. Krautmann and Sander (1999) find similar results by using instrumental variables. But when they exploit random

assignment to teachers, Braga, Paccagnella, and Pellizzari (2014) and Carrell and West (2010) show that good teachers get bad evaluations. Trying to disentangle grades and learning, Beleche, Fairris, and Marks (2012) find a positive correlation between their measure of student learning and course evaluations, but Weinberg, Hashimoto, and Fleisher (2009) find that evaluations are positively correlated to grades but unrelated to learning.

In this paper, we try to identify the effect of grade on the evaluation of teaching. Once any simultaneity or selection biases are ruled out, at least two reasons are invoked in the literature to explain such a link. Students may infer course quality from received grades this is the "attribution" theory of Greenwald and Gillmore, 1997. Or students may give good evaluations as a reward for good grades or easy exam ("leniency" theory). Using an original data set of individual grades and evaluations from a French higher education institution, we ask whether students do account for their grade when evaluating the quality of teaching. We contribute to the literature in two ways. First, to our knowledge, this is the first study to analyze the causal link between grades and SET in the French context.³ Second, compared to other data sets, we know the exact date when students evaluate each course, so that we are able to analyze the dynamics of evaluations over time. We also use the traditional methods proposed in the literature, i.e. teacher's fixed effects and instrumental variables. We find that students give teachers higher evaluations when they get higher grades, but the size of the effect is small. We also find that students use contemporaneous information into account when evaluating teaching. They give higher evaluations after the exam, and lower evaluations after getting their grades.

The remaining is organized as follows. We present the data in Section 3.1. Section 3.2 gives some descriptive statistics. Section 3.3 presents the empirical strategy and gives the results. We conclude in Section 3.4.

3.1 The data

We use data from a French "*grande école*", which provides a three-year program in the fields of statistics, economics, finance and actuarial science. The first two years consists in the same basic training for all students. In third year, students specialize into one particular track.

³We are aware of only one study, Boring (2015), using French data to analyze gender biases in student evaluations of teachers.

Since the 2004-2005 school year, each student has been asked to evaluate the courses they attend. Students fill an online form, which is the same for every course. Evaluation is completely anonymous. The evaluation form consists of seven questions about the course (interest, difficulty, student assiduousness, teacher's pedagogy, formalization, progression and quality of course material), and five about tutorial classes when they exist (difficulty, student assiduousness, teaching assistant's pedagogy, number of tutorials, consistency with the course). Questions regarding the course material, the number of tutorials and the consistency with the course were removed from this analysis. For each question, students have to choose between three to four ranked categories. The detail of questions and categories is given in Table 3.4 of the appendix. Students can also write free comments.

In France, students' evaluation of teaching has been introduced only recently and is not commonly used to evaluate teachers. For now, recruitment and promotion procedures do not depend on SET at all. Teachers are usually civil servants, and they are recruited on the basis of a national competitive examination that does not take their previous evaluations into account. Teacher's pay is based on a national salary scale, and does not depend on SET. However, each higher education institution administration may use SET to organize teaching. In the school we study, the administration may change teacher in a particular course if students evaluations are too bad for several consecutive years. Note that the use of SET by the administration may differ for teachers and for teaching assistants. Because teaching assistants are not permanent staff of the school, the administration may more easily choose not to keep them if their evaluations are too bad.

We use three anonymous data sets. First we observe every courses given in the school from the 2004-2005 to the 2010-2011 school year. We excluded from our analysis some types of courses: courses without notation, language courses, and collective projects, for which there is no related teacher. Furthermore, the form is not suited for courses taught by several teachers, so we excluded them. We also excluded small classes (less than 10 students registered), or courses which were given only one or two school years. In the end, we observe 97 courses, from 2004-2005 to 2010-2011, that is 485 observations. This is an unbalanced panel: not all courses in the sample are given every year, but they are all given at least two years. Among them, 17 are given in the first year of the school program, 27 in second year and 53 in third year. This reflects increasing specialization over the program. 37 courses have tutorial classes, 60 have not. Overall, we observe 128 individual teachers and

291 teaching assistants. Every teacher gives between 1 and 5 courses per school year and they give 1.3 course a year on average. Every teaching assistant gives tutorial classes for 1 to 4 courses each year, with 1.5 on average. Courses are split into themes: 21 in economics, 12 in social sciences, 21 in finance and actuarial science, 16 in mathematics and computer science and 27 in statistics and econometrics.

Second, we observe individual evaluations of each course. This corresponds to 17 560 individual observations. Note that the data are individual but anonymous; we cannot link different evaluations made by the same student. We observe, though, some individual characteristics, namely the way students entered the school (competitive examination or direct admission)⁴, and the date when they evaluate courses. We also observe course-specific characteristics: whether it takes place in year 1, 2, or 3, the major, one identifier for the course and one for the teaching assistant. The number of very bad evaluations (category 1) is small (see Table 3.5 in the appendix). On average, 8% of students find that the teacher's pedagogy is bad. 19% rate it as fair, 49% as good and 24% as very good.

Third, we have a data set of individual grades for each course, containing 24 198 individual observations. Again, this data set is anonymous; we cannot link different grades obtained by the same student for different courses. Students may be graded through two different types of evaluations, depending on the course. First, they may have to take a written exam at the end of the semester. The exam is usually prepared and corrected by the teacher, sometimes with the help of teaching assistants. In the course panel, 359 observations out of 485 are evaluated through a written exam. Second, students may get a continuous assessment grade, by handing in a homework, or a dissertation, or by attending tutorial classes. If there is no written exam, then the final grade is given by the teacher, on the basis of a continuous assessment evaluation. This is the case for 126 observations out of 485. If there is a written exam and a continuous assessment (112 observations out of 485), then the final grade is a weighted average of both grades. In this case, the continuous assessment is usually made by the teaching assistant, whereas the exam is graded by the teacher. For each individual observation in the grades data set, we know the individual written exam grade and continuous assessment grade, as well as the grade after retake, if any. We observe student's way of admission, one identifier for the

⁴Students can enter the school through different ways and at different stages: students are admitted either in the first year through a competitive examination (mathematics or economics major), or directly in second or third year on the basis of academic qualification.

teacher and one for the teaching assistant. Grades are different across courses; mean grades in social sciences courses are higher than in courses with a more technical content (Table 3.6 in the appendix).

The diploma delivered at the end of the school is the same whatever the grades students get, and is a very good signal on the French job market. In that sense, grades may matter less than in other higher education institutions, where students may usually get distinctions depending on their grades. However, the delivering of the school's diploma is subject to some rules regarding students' grades. The diploma is delivered at the end of Year 3 of the school, but in order to pass each year, students have to get an average of 12 over 20 during the year and to get at least a 6 in some core subjects. Students may repeat one year if they fail to meet these rules, but they cannot repeat more than one year during the school. Furthermore, students are allowed to, and often do, attend another Master during Year 3 of the school. These students usually try and attend highly selective Masters, and thus need to get very good grades during the first two years of the school. Similarly, students willing to continue in a Ph.D. track may need good grades to get funding.

Because data sets are anonymous, we cannot link individual grades and evaluations. Thus our analysis is made at the aggregate teacher-year-subject level, which is in line with what is done in the literature. It would be possible to work at a more detailed level, by using the teaching assistant identifier. The number of observations would be larger. Moreover, as tutorial classes are formed almost randomly⁵, there would be less worry concerning a potential selection of students within classes. However, contrary to the teacher's name, the teaching assistant's name is self-declared by students and is of poor quality. More precisely, students have to select the name of the teaching assistant from a drop-down menu, and the first name appearing in the menu systematically has much more evaluations than the others. We tried and correct this by deleting observations such that the teaching assistant has more than twice as much evaluations than she should have if groups were of equal size. So, partly because we lose a lot of (potentially not random) observations, we will use the observations aggregated at the teaching assistant-year-subject level with caution.

At the end of each semester, students can evaluate the courses they attended whenever they want by filling out the online form. The exact date when each

⁵Depending on the school year, students are allocated to tutorial classes using alphabetical order, or a more random allocation.

evaluation is made is observed. Whatever the school year, there is a clear mode in the number of evaluations over time (see Figure 3.2 in the appendix). We assume that this mode corresponds to the date when grades are released. For each course, we are thus able to define the (unobserved) date when grades were released as the mode of the dates of individual evaluations. Because we know the date of the final written exam for each course, we are then able to study the dynamics of evaluations, according to the level of information students have on their expected grade.

Student evaluation of teaching was optional up to 2007-2008 and could be made after the final grades were released. From the 2008-2009 school year, it became compulsory and grades are now obtained only after having filled out the form. Thus, students who completed the evaluation before 2008 are likely to have specific characteristics. Students response rate is not observed, but, linking evaluations and grades at the teacher-year-subject level, we can observe the number of evaluations over the number of grades for each course. This ratio is 45% on average before 2008. After 2008, it is 100% on average. Moreover, although we do not observe any large differences in mean grades before and after 2008 (Table 3.6 in the appendix), mean evaluations significantly differ before and after 2008 (see Table 3.7 in the appendix). Before 2008, students report to be more assiduous on average. They rate the teacher's pedagogy to be lower, and they rate the speed of progression higher. They also write comments more often. Students evaluating before 2008 are more often in Year 2, or in Year 3 of the school program, and they entered more often the school through direct admission. In the following estimations, we will thus be careful about controlling for before or after 2008.

3.2 Descriptive statistics

Correlations between the different dimensions of teaching evaluated by students (Table 3.8 in the appendix) highlight two groups of variables: teacher's pedagogy, interest for the subject, and assiduousness are positively correlated with one another and negatively correlated with the difficulty of the course, formalization and speed of progression. Mean grades are positively correlated with the dimensions of the first group and negatively correlated with the variables of the second group. The same results emerge from a principal component analysis (Figure 3.4 in the appendix), where the first two axis explain around 70 % of the total variance. The two groups appear clearly on either side of the first axis in the aggregate level analysis.

Figure 3.3 in the appendix confirms the direction of correlations between mean grades and evaluations. One more point in mean grade corresponds to 0.04 point more in the mean evaluation of teacher's pedagogy (and to 0.03 point more in the evaluation of the interest for the subject and in assiduousness). When the course progresses too fast, when it is too formalized or when it is too difficult, mean grade is smaller. Those effects are as expected, but they are quantitatively small, about one tenth of a standard error. This is partly due to the fact that evaluations are categorical ordered variables.

Unsurprisingly, variance analysis in Table 3.9 of the appendix shows that most part of the variance in evaluations and grades comes from between students variation and that inter-course variance is small.

In the following, we will consider teacher's pedagogy as the dependent variable. The teacher's pedagogy item seems to us as the best proxy of how students evaluate teaching quality. And, in practice, this is the main dimension that the school administration uses to assess the quality of teaching. Potential determinants of by-course mean evaluation of teacher's pedagogy are presented in Table 3.1. The first column gives the ordinary least squares estimates of the regression of mean teacher's pedagogy on course characteristics (observations are clustered at the course level). Students' mean evaluation of teacher's pedagogy is negatively correlated with the number of students in the class. This is in line with the literature considering that the teacher has less time to devote to each student when the class is bigger. The fact that the course is evaluated through a written exam is positively associated with the mean evaluation of teacher's pedagogy. The domain of the course (human sciences, finance and actuarial sciences, mathematics and computer sciences, or statistics and econometrics; economics is the reference) is not significantly correlated with the mean evaluation of teacher's pedagogy. As already mentioned, mean evaluation is higher after 2008 than before. The second column of Table 3.1 adds observable characteristics of the teacher as regressors. As expected, no teaching experience, as measured by a dummy for the first year of teaching the course, is negatively correlated with the mean evaluation of the level of pedagogy. Whether the teacher is male or female is not significantly correlated to the mean evaluation he or she gets (contrary to the results of Boring, 2015). Columns 3 and 4 add mean students' grades as covariates. The mean final grade is positively associated to the mean evaluation of teacher's pedagogy, which is actually driven by a positive correlation with the mean exam grade. Mean continuous assessment grade, when there is also a written exam,

is positively but not significantly associated with the mean evaluation. When there is no written exam, i.e. when the continuous assessment grade is the only way in which students are graded, then the correlation is negative, although not significant. Column 5 gives mean students' evaluations of their interest for the subject as an additional regressor. As expected, the correlation is significantly positive. Column 6 shows that the mean evaluation of teacher's pedagogy the preceding year is positively correlated with the current year mean evaluation.

In the following, we will further analyse the relationship between grades and evaluation of teacher's pedagogy.

Table 3.1 – OLS estimations of potential determinants of teacher’s pedagogy mean evaluation

	Teacher’s pedagogy					
	(1)	(2)	(3)	(4)	(5)	(6)
Nbr students	-0.003** (0.00)	-0.003** (0.00)	-0.003** (0.00)	-0.003** (0.00)	-0.001 (0.00)	-0.001 (0.00)
Written exam	0.100* (0.06)	0.090 (0.06)	0.254*** (0.08)	-1.198 (0.75)	-1.607** (0.67)	-1.105 (0.67)
Human sciences	0.041 (0.13)	0.024 (0.14)	0.009 (0.13)	-0.008 (0.12)	-0.098 (0.13)	-0.166 (0.11)
Finance	-0.044 (0.13)	-0.036 (0.13)	-0.024 (0.13)	-0.025 (0.13)	-0.114 (0.12)	-0.105 (0.10)
Maths, computer	0.012 (0.12)	0.011 (0.12)	-0.004 (0.12)	0.003 (0.12)	0.041 (0.11)	-0.022 (0.09)
Stats, Econometrics	0.006 (0.13)	0.015 (0.13)	-0.002 (0.13)	0.025 (0.13)	0.024 (0.11)	0.030 (0.08)
After 2008	0.124* (0.06)	0.130 (0.08)	0.121 (0.08)	0.117 (0.08)	0.103 (0.06)	0.135** (0.05)
Intercept	2.828*** (0.17)	2.834*** (0.19)	1.762*** (0.32)	3.124*** (0.71)	1.492** (0.67)	0.636 (0.72)
First year teaching		-0.188*** (0.07)	-0.184*** (0.06)	-0.174*** (0.06)	-0.133** (0.07)	-0.072 (0.08)
Male prof		0.047 (0.09)	0.047 (0.09)	0.041 (0.09)	-0.011 (0.11)	0.001 (0.11)
Mean final grade			0.073*** (0.02)			
Mean exam grade				0.081*** (0.02)	0.056*** (0.02)	0.050*** (0.01)
Mean cc with exam				0.007 (0.01)	-0.005 (0.01)	-0.004 (0.01)
Mean cc without exam				-0.020 (0.05)	-0.065 (0.04)	-0.034 (0.04)
Mean interest					0.774*** (0.08)	0.654*** (0.09)
L.Mean pedagogy						0.270*** (0.06)
Dummies school grade	Yes	Yes	Yes	Yes	Yes	Yes
Dummy direct admission	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.07	0.10	0.13	0.14	0.39	0.48
Nb obs	485	485	485	485	485	373
Nb clusters	97	97	97	97	97	97

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). The unit of observation is a triplet teacher-course-year. Observations are weighted by the number of students in the unit. Standard errors in brackets are clustered at the course level.

Lecture: On average in the sample, the teacher’s pedagogy is rated 2.828 out of 4 before 2008 and 2.952 (=2.828+0.124) after 2008. This difference is significant at the 10% significance level.

3.3 Empirical strategy and results

3.3.1 Empirical strategy

When analyzing the effect of grade on the evaluation of teaching, we first need to address the simultaneity issue. Better teaching leads to better grades and better evaluations at the same time. In order to better understand the link between evaluations and grades, let us consider the following model:

$$Eval_{icjt} = \alpha_0 + \alpha_1 Grade_{icjt}^e + \alpha_2 X_{it}^1 + \alpha_3 Y_{ct}^1 + q_{cj} + u_{icjt} \quad (3.1)$$

where subscript i corresponds to the student, c corresponds to the course, j corresponds to the teacher, and t corresponds to the school year. $Eval_{icjt}$ is thus student i 's evaluation of teacher j 's pedagogy for course c in school year t . $Grade_{icjt}^e$ is the grade student i expects to receive from teacher j for course c in school year t . X_{it}^1 and Y_{ct}^1 are student and course specific observed characteristics respectively. q_{cj} is the unobserved pedagogical qualities of teacher j teaching course c . Error terms u_{icjt} are self-correlated within course. If student evaluation was purely objective (and if they had complete information on teaching quality), α_1 would be 0, i.e. students' evaluations would depend only on observed and unobserved course characteristics, and not on their expected grade. In other words, students would not react to their grade when evaluating the course. We want to know whether this is true or not. α_1 is thus our parameter of interest.

The difficulty comes from the fact that student i 's expected grade is also explained by observed and unobserved individual and course characteristics:

$$Grade_{icjt}^e = \gamma_0 + \gamma_1 X_{it}^2 + \gamma_2 Y_{ct}^2 + q_{cj} + \gamma_i + v_{icjt}$$

with γ_i an unobserved student fixed effect corresponding to student i 's individual ability. We assume that u_{ij} and v_{ij} are not correlated.

This equation shows that we also need to take into account possible endogeneity issues due to the fact that teacher's quality and student's ability are unobserved and potentially correlated with observed variables. More formally, there is a correlation between $Grade_{icjt}^e$ and the error term $q_{cj} + u_{icjt}$ in equation (3.1), due to student's fixed effect, teacher's fixed effect, and possible student's selection into courses.⁶

⁶ $E[Grade_{icjt}^e(q_{cj} + u_{icjt})] = E[(\gamma_0 + \gamma_1 X_{it}^2 + \gamma_2 Y_{ct}^2 + q_{cj} + \gamma_i + v_{icjt})(q_{cj} + u_{icjt})] = E(q_{cj}^2) + E(\gamma_i u_{icjt})$. The correlation is the sum of a teacher's quality effect and a students' selection into course effect.

Again, note that, because data are anonymous, we cannot link individual grades and individual evaluations. In the following, we will then consider mean evaluations and mean grades by teacher-course-year. We will note $\bar{x}_{cjt} = \frac{1}{N_{cjt}} \sum_{i=1}^{N_{cjt}} x_{icjt}$, where N_{cjt} is the number of students enrolled in course c , with teacher j , in school year t . We then consider the pseudo-panel model aggregated at the teacher-course-year level:

$$\overline{Eval}_{cjt} = \alpha_0 + \alpha_1 \overline{Grade^e}_{cjt} + \alpha_2 \overline{X^1}_{cjt} + \alpha_3 Y_{ct}^1 + q_{cj} + \bar{u}_{cjt} \quad (3.2)$$

Note that the errors are inherently heteroscedastic in this model. In order to get efficient estimates, regressions are weighted by the number of students per course.

Let us consider that students know their grade or at least have a good expectation of it, so that their expectations are correct on average, i.e. $\overline{Grade^e}_{cjt}$ is replaced by \overline{Grade}_{cjt} in equation (3.2).

Our first identification strategy consists in estimating equation (3.2) with teacher-course fixed effects. Identification is then based on variations from one year to another in a given course taught by a given teacher. This is manageable only for teacher-course pairs that we observe several times. Moreover, for grades not to be correlated with the error term, we have to assume that i) $\bar{\gamma}_{cjt}$ is independent from \bar{u}_{cjt} , i.e. course average students' ability is independent from evaluations unobserved characteristics, and that ii) \bar{v}_{cjt} is independent from \bar{u}_{cjt} , i.e. there is no common idiosyncratic shock affecting both grades and evaluations. Assumption i) is valid if course average students' ability is constant over years. This seems realistic, at least for compulsory courses, or if course characteristics are properly controlled for. Assumption ii) seems realistic in general. It would not be valid if, for instance, the course took place at some very inconvenient time of the day.

Our second identification strategy consists in estimating equation (3.2) with two stage least squares, by instrumenting mean grade \overline{Grade}_{cjt} . We thus need to choose instruments that would explain mean grade but not evaluations directly. A classical choice in the literature is to use mean grade the preceding year, that we will note $\overline{Grade}_{cj,t-1}$ (provided that the teacher is the same in t and $t-1$). However, in our setting, the exclusion restriction is not very credible because lagged mean grade contains the unobserved quality of teacher. Thus, we propose to use the lagged mean grade only when the teacher changes between $t-1$ and t . In this case, mean grade the preceding year is more likely to be uncorrelated with \bar{u}_{cjt} . We will also use the proportion of retake as instrumental variable. This is a good predictor of mean grade, that should not be correlated to unobserved characteristics of mean evaluation.

3.3.2 Results

Table 3.2 presents the estimation results using OLS, teacher's fixed effects, and 2SLS. The first outcome we consider is the mean evaluation of teacher's pedagogy by students enrolled in the course. In order to further analyse the distribution of evaluations, two other outcomes are considered, namely the proportion of students giving very good evaluations, and the proportion of very bad evaluations. The explanatory variables of interest are the course mean written exam grade, and the course mean continuous assessment grade, interacted with a dummy which equals one if there is a written exam and zero otherwise.

The null hypothesis we want to test is that students do take their grades into account when evaluating teacher's pedagogy. If this is true, then the coefficient of mean exam grade should be significant, even after correcting for endogeneity. The corollary is that the mean evaluation of teacher's pedagogy should not depend on the mean continuous assessment grade when there is a written exam, since, in this case, continuous assessment is made by teaching assistants. On the contrary, when there is no written exam, the mean evaluation of teacher's pedagogy should depend on the mean continuous assessment grade.

Regressions are controlled for course characteristics and teacher characteristics. Course characteristics are the domain (economics, human sciences, etc.), whether the course is given in year 1, 2 or 3 of the school program, and whether this is a catch-up course for students entering directly in year 2 or 3. Teacher characteristics are a dummy for males, and a dummy for the first year of teaching. Because we are concerned by a potential selection issue in evaluations before 2008, a dummy for observations after 2008 is also added as a covariate.

The first column of Table 3.2 shows a significantly positive correlation between mean teacher's pedagogy evaluation and the mean grades which are given by the teacher. When there is a written exam, mean continuous assessment grade is negatively, though not significantly, correlated with mean teacher's pedagogy evaluation. These naive estimates are thus in line with the hypothesis that students do take their grades into account when evaluating teaching. More precisely, a one point increase in mean exam grade would correspond to a 0.077 increase in teachers' pedagogy mean evaluation, going from 1.803 on average to 1.880. However, this estimate does not take the endogeneity of mean grades into account. When endogeneity is controlled for, using teacher-course fixed effects in column 2, and using instrumental variables in column 3, the coefficient of mean exam grade remains

positive and significant. The size of the effect is a bit larger when using instrumental variables. The coefficient associated to mean continuous assessment grade without a written exam is larger in the fixed effect specification, but not significant in the instrumental variables one. The effect of mean continuous assessment grade, when not given by the teacher, remains not significantly different from zero. The positive effect of grade on mean evaluation seems to be driven by a smaller proportion of very bad evaluations. A one point increase in mean continuous assessment grade, when given by the teacher, corresponds to a 0.95 percentage point decrease in the proportion of very bad evaluations that the teacher gets. The effects of grades on the proportion of very good evaluations of teacher's pedagogy are not significantly different from zero, in either specifications (except for continuous assessment grade with exam in the OLS specification).

In the 2SLS specification, mean grade the preceding year, and percentage of retake seem to be valid instruments for current mean grade. According to Table 3.3, the coefficients of lagged mean grades are highly correlated to corresponding current mean grades in the first stage estimations. When there is a written exam, the proportion of retake is also a significant determinant of mean grade. Furthermore, the test for joint significance ("F weak" statistic) rejects the weak instruments hypothesis.

Again, the exclusion restriction is unlikely to hold for lagged mean grade. Tables 3.10 and 3.11 in the appendix thus replicate the preceding results, but the two stage least squares are estimated only on observations for which the teacher changed between $t - 1$ and t . First stage estimates are of course not as significant as before, and the F statistics are much lower, although higher than 10. In the second stage, the number of observations drop from 314 to 59, and the intercept is out of the support. The sign of the coefficients however is as expected, and an increase in the mean grades given by the teacher significantly decreases the proportion of very bad evaluations.

To confirm these results, we tried and estimated the same models at the teaching assistant-course-year level. Symmetrically, we should find that mean teaching assistant's pedagogy evaluation does not depend on the mean grade given by the teacher, but depends on the mean continuous assessment grade given by the teaching assistant (i.e. when there is a written exam). Results are given in Table 3.12 in the appendix. The first column of Table 3.12 shows a positive, but hardly significant, correlation between mean continuous assessment grade given by the

teaching assistant and his or her pedagogy mean evaluation. This effect is not significantly different from zero, however, in the teaching assistant-course fixed effects and instrumental variables specifications. Contrary to what was expected, the mean exam grade is negatively associated to teaching assistant's pedagogy mean evaluation in the IV setting. Moreover, an increase in mean exam grade significantly decreases the proportion of very good evaluations the teaching assistant gets. When there is no written exam, an increase in mean continuous assessment grade gives about the same result and decreases teaching assistant's pedagogy mean evaluation, although the proportion of very good evaluations increases. As mentioned earlier, we are not very confident in the quality of the data, so that these results should be taken with caution.

To sum up, our results suggest a positive relationship between students evaluation of teacher's pedagogy and grades which are given by the teacher him or herself. We find no significant effect of grades given by the teaching assistant. This is in line with the literature. We cannot however distinguish the channels of such an effect. Do students reward (respectively punish) teachers for lenient (respectively severe) grading? Or do they attribute a good grade to a good teaching? To try and learn more about that, we propose to study the dynamics of evaluations over time.

Table 3.2 – Estimation of the effect of mean grade on teacher’s pedagogy mean evaluation

Teacher’s pedagogy	mean			%very good			%very bad		
	OLS	FE	IV	OLS	FE	IV	OLS	FE	IV
Mean exam grade	0.077*** (0.022)	0.072* (0.037)	0.086* (0.047)	0.124 (0.236)	0.447 (0.326)	-0.576 (0.463)	-0.762** (0.314)	-0.552 (0.424)	-0.645 (0.628)
Mean cc without exam	0.059*** (0.018)	0.094** (0.036)	0.061 (0.039)	0.093 (0.204)	0.293 (0.339)	-0.521 (0.384)	-0.767*** (0.286)	-0.951** (0.438)	-0.657 (0.531)
Mean cc with exam	-0.001 (0.008)	0.005 (0.012)	-0.002 (0.010)	0.219** (0.094)	-0.075 (0.161)	0.147 (0.106)	-0.086 (0.085)	0.157 (0.115)	-0.022 (0.114)
Intercept	1.803*** (0.276)	1.856*** (0.464)	1.815*** (0.561)	15.350*** (3.551)	15.749*** (3.896)	23.506*** (5.737)	23.670*** (4.229)	18.406*** (5.148)	21.577*** (7.913)
Course characteristics	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
After 2008 dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.11	0.77	0.08	0.20	0.73	0.21	0.12	0.72	0.13
Nb obs	485	485	314	485	485	314	485	485	314
Nb clusters	97	97	96	97	97	96	97	97	96

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). The unit of observation is a triplet teacher-course-year. Observations are weighted by the number of students in the unit. Standard errors in brackets are clustered at the course level. Course characteristics = dummies for type of subject, dummies for year 1, 2 or 3 of the school program, dummy for direct admission catch-up course. Teacher characteristics = dummy for male, dummy for first year of teaching. The FE specification contains teacher-course fixed effects. In the IV specification, instrumental variables = lagged mean exam grade, lagged mean continuous assessment grade without exam, lagged mean continuous assessment grade with exam, and proportion of students retaking the exam.

Lecture: On average in the sample, 23.670% of students in a class rate the teacher’s pedagogy as very bad. A one point increase in mean exam grade is associated with a 0.762 decrease in this proportion. This coefficient is significant at the 5% significance level.

Table 3.3 – First stages

First stage	Mean exam	Mean cc without exam	Mean cc with exam
Lag mean exam	0.377*** (0.087)	0.115* (0.068)	0.007 (0.107)
Lag mean cc without exam	-0.479*** (0.060)	0.995*** (0.036)	0.037 (0.097)
Lag mean cc with exam	-0.014 (0.016)	0.008 (0.010)	0.867*** (0.058)
% retake	-0.049*** (0.013)	-0.007 (0.010)	0.060* (0.032)
Intercept	6.755*** (1.013)	-0.793 (0.727)	1.064 (1.367)
Course characteristics	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes
After 2008 dummy	Yes	Yes	Yes
R2	0.86	0.92	0.92
Fstat 1st stage	281	507	242
Nb obs	314	314	314
Nb clusters	96	96	96

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). The unit of observation is a triplet teacher-course-year. Observations are weighted by the number of students in the unit. Standard errors in brackets are clustered at the course level. Course characteristics = dummies for type of subject, dummies for year 1, 2 or 3 of the school program, dummy for direct admission course. Teacher characteristics = dummy for male, dummy for first year of teaching.

Lecture: A one point increase in mean exam grade the preceding year is associated with a 0.377 increase in current mean exam grade. This coefficient is significant at the 1% significance level.

3.3.3 The dynamics of evaluations over time

Students have a different set of information, both on the quality of the course, and on their own achievement, depending on when they evaluate. More precisely, we are going to have a closer look at the dynamics of students evaluations with respect to the date when they take the written exam, and the date when they get their grades.

To study the relationship between students evaluation of teacher's pedagogy and the date of evaluation, we propose to use a very flexible model. More precisely, we use a generalised additive model, assuming that the expected value of the evaluation

variable is an unknown function of the date of evaluation, in an additive relationship. The model is of the form:

$$g[\mathbb{E}(Eval_{icjt}|dist, u)] = \alpha + f(dist_{icjt}) + u_{ct} \quad (3.3)$$

As before, $Eval_{icjt}$ measures student i teacher's pedagogy evaluation.⁷ Individual dates of evaluations are centered in the exam date, so that $dist_{icjt}$ is the individual distance to exam date, in days. In a second model, we also consider the distance to grade release.⁸ The link function g is assumed to be a normal distribution and f is estimated using splines (of degree that is chosen using a generalised cross-validation method). Course-year fixed effects u_{ct} are added to control for differences in evaluations across courses and year.

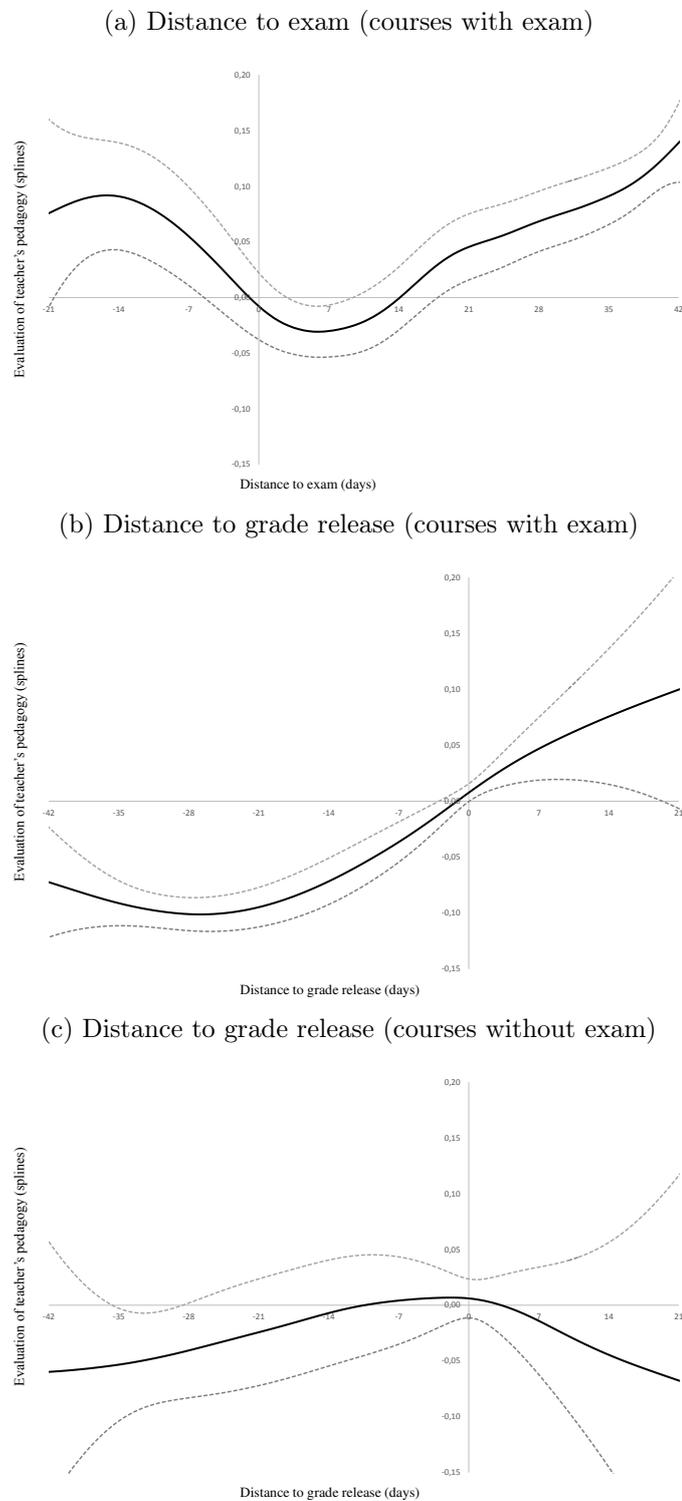
Figure 3.1 gives a graphical representation of the estimation of function f , separately for distance to exam, distance to grade release, for courses with and without exams. First, note that there is no significant relationship between evaluations and distance to grade release, for courses which are not graded through a written exam (Graph c). This result implies that students do not use (or do not react to) the information given by their grades to evaluate teacher's pedagogy, in the case when the course is graded through a continuous assessment. Moreover, this suggests no selection effect with respect to the date of evaluation. In other words, students evaluating sooner do not rate teachers higher or lower. On the contrary, teachers giving a written exam are evaluated differently according to the date of evaluation (Graphs a and b). More precisely, Graph a shows a small but significant decrease in \hat{f} after the exam, meaning that students rate teachers lower after taking the exam (for about a week). This could be due to an effect of deception, or a way of punishing teachers. Note, however, that the decrease begins before the exam, and could thus hide selection effects, i.e. content students may evaluate sooner. After grades are released (Graph b), \hat{f} significantly increases, meaning that students rate teacher's pedagogy higher on average after getting their grades. Again, this could suggest either a reward, or an attribution of good grade to good teaching. Although we cannot clearly separate those two effects, these results seem very important, because they highlight that the way students evaluate teaching highly depends on

⁷We used a dummy variable for very good (respectively very bad) evaluations as an alternative outcome. The results are in line with our preceding results; very good evaluations do not depend on the distance to grade release.

⁸A model using both variables (distance to exam, and distance to grade release) together was also estimated. The results are similar.

the contemporaneous information they get.

Figure 3.1 – Estimation of the relationship between evaluations and date



Note: Each graph represents the estimation of function f through a generalized additive model, with course-year fixed effects. Function f is estimated using splines of degree 6 for distance to exam, and 3 for distance to grade release. Dashed lines give confidence intervals.

3.4 Conclusion

This chapter proposes to evaluate students' objectivity in their evaluation of teaching. We use a unique data base in the context of a French higher education institution. First, we analyse the relationship between grades and evaluations. Using teacher-course fixed effects and instrumental variables, we find that students do take their grades into account when they evaluate the pedagogy of a teacher. The relationship is positive, suggesting either that students reward teachers for good grades, or that they attribute a good grade to a good teaching. The size of the effect is relatively small and driven by a decrease in the proportion of discontent students; a one point increase in by-course mean exam grade corresponds to less than a 1 percentage point decrease in the proportion of students giving very bad evaluations. Second, we try and study the timing of evaluations, using information on the exact date when students evaluate teaching. We find that students use contemporaneous information when they evaluate the teacher's pedagogy. They give lower evaluations after taking the exam, and higher evaluations after getting their grades. When students are graded through a continuous assessment, however, they do not seem to react to the release of their grades.

Our results are based on data from a particular French higher education institution. They may not be representative of every higher education institutions, nor even of every French universities. The institution studied here is a very selective school, and delivers a degree that is a good signal on the French labor market. Once they have been selected to enter the school, students may not care for the grades they get as much as other students. This may partly explain why we find relatively small effects. Although we cannot compare our results to other French institutions, we can thus speculate that students grades would bias their evaluation of teaching even more in other institutions. Moreover, teachers' recruitment procedures in this school is different than in other institutions, so that the use of SET by the administration is presumably not the same. It may be that SET matters more for the organization of teaching in this particular school than in other French higher education institutions.

These results lead to two conclusions. First, they confirm that evaluations may be distorted by teachers trying to buy good evaluations, or by students trying to get information about the quality of the course through the exam or through the grades they get. Second, our results highlight that students have difficulties evaluating teaching and use available information to do it. A solution may be to make students evaluate all together at a single date in time. This may homogenize the information

they get, but this would not prevent students from using the contemporaneous information they have at that date.

Our purpose is not to recommend against the use of SET. On the contrary, it has proven to be a relevant measure of (at least some aspects of) teaching quality. After all, who but the persons who attend the course may best evaluate the quality of teaching? However, institutions should be aware of potential distortions both teachers and students are likely to create. In order to obtain an objective measure of teaching quality, the administration may want to also rely on other ways to evaluate teaching.

Appendix

Table 3.4 – Description of the evaluation form

Questions	Answers			
	1	2	3	4
Interest for the subject	<input type="checkbox"/> Not interesting	<input type="checkbox"/> Moderately interesting	<input type="checkbox"/> Interesting	<input type="checkbox"/> Very interesting
Difficulty of the course	<input type="checkbox"/> Easy	<input type="checkbox"/> Moderate	<input type="checkbox"/> Difficult	<input type="checkbox"/> Very difficult
Assiduousness in courses	<input type="checkbox"/> Less than half	<input type="checkbox"/> About half	<input type="checkbox"/> All courses or almost	
Teacher's pedagogy	<input type="checkbox"/> Bad	<input type="checkbox"/> Fair	<input type="checkbox"/> Good	<input type="checkbox"/> Very good
Level of formalization	<input type="checkbox"/> Too low	<input type="checkbox"/> Adequate	<input type="checkbox"/> Too high	
Speed of progression	<input type="checkbox"/> Too slow	<input type="checkbox"/> Adequate	<input type="checkbox"/> Too fast	
Difficulty of the tutorials	<input type="checkbox"/> Easy	<input type="checkbox"/> Moderate	<input type="checkbox"/> Difficult	<input type="checkbox"/> Very difficult
Assiduousness in tutorials	<input type="checkbox"/> Less than half	<input type="checkbox"/> About half	<input type="checkbox"/> All tutorials or almost	
Teaching assistant's pedagogy	<input type="checkbox"/> Bad	<input type="checkbox"/> Fair	<input type="checkbox"/> Good	<input type="checkbox"/> Very good

Lecture: To the question concerning their interest for the subject, students could answer that they find it not interesting, interesting, moderately interesting, or very interesting. The variable is coded 1, 2, 3 or 4 respectively.

Table 3.5 – Distribution of evaluations in the sample

	Interest	Difficulty	Assiduousness	Pedagogy	Formalization	Progression
	%	%	%	%	%	%
1	5	3	16	8	7	4
2	13	37	27	19	78	76
3	49	45	57	49	15	20
4	33	15		24		
Total	100	100	100	100	100	100
Observations	17337	17260	17295	17254	17082	17016

Lecture: On average, 5% of students find that the subject of the course is not interesting and 33% find it very interesting. See table 3.4 for the meaning of each category.

Table 3.6 – Main characteristics of exam grades in the sample

	Before 2008		After 2008		Test (pvalue)
	N	(mean/sd)	N	(mean/sd)	
Grades					
<i>Economics</i>	1,966	12.12 0.31	2,760	11.47 0.47	0.106
<i>Human sciences</i>	817	13.48 0.47	1,343	13.10 0.45	0.293
<i>Finance and actuarial sciences</i>	3,387	12.08 0.31	2,441	12.58 0.34	0.155
<i>Mathematics and computer sciences</i>	1,948	11.81 0.87	2,473	11.84 0.38	0.974
<i>Statistics and econometrics</i>	3,479	11.48 0.42	3,584	12.02 0.38	0.060

Note: Standard errors are clustered at the course level.

Lecture: On average, students get 12.12 out of 20 at the exam before 2008 and 11.47 after 2008, for courses of the major Economics. The difference is not significantly different from zero.

Table 3.7 – Main characteristics of evaluations in the sample

	Before 2008		After 2008		Test (pvalue)
	N	(mean/sd)	N	(mean/sd)	
Evaluations					
<i>Interest for the subject</i>	4,735	3.07	12,602	3.09	0.699
		0.03		0.03	
<i>Difficulty</i>	4,682	2.71	12,578	2.73	0.698
		0.05		0.04	
<i>Assiduousness</i>	4,722	2.49	12,573	2.39	0.001
		0.03		0.02	
<i>Pedagogy</i>	4,627	2.82	12,627	2.93	0.084
		0.07		0.04	
<i>Formalization</i>	4,553	2.10	12,529	2.07	0.174
		0.02		0.02	
<i>Speed of progression</i>	4,531	2.19	12,485	2.13	0.027
		0.03		0.02	
<i>Comments=1</i>	4,923	0.46	12,637	0.23	0.000
		0.01		0.01	
School year					
<i>1st=1</i>	4,027	0.14	12,626	0.27	0.002
		0.04		0.06	
<i>2nd Eco=1</i>	4,027	0.19	12,626	0.16	0.256
		0.04		0.03	
<i>2nd Fin=1</i>	4,027	0.13	12,626	0.20	0.000
		0.02		0.03	
<i>2nd Stat=1</i>	4,027	0.14	12,626	0.08	0.003
		0.02		0.01	
<i>3rd=1</i>	4,027	0.41	12,626	0.29	0.017
		0.07		0.05	
Admission					
<i>Direct=1</i>	3,900	0.38	12,617	0.32	0.003
		0.02		0.03	
<i>Eco=1</i>	3,900	0.25	12,617	0.27	0.150
		0.02		0.02	
<i>Math=1</i>	3,900	0.37	12,617	0.41	0.015
		0.02		0.02	

Note: Standard errors are clustered at the course level.

Lecture: On average, students rate 3.07 out of 4 the interest for the subject before 2008 and 3.09 after. The difference is not significantly different from zero.

Table 3.8 – Correlations between mean evaluations and mean grades

	Interest	Difficulty	Assiduousness	Pedagogy	Formalization	Progression	Grade
Interest	1.00						
Difficulty	-0.01	1.00					
Assiduousness	0.44***	-0.05	1.00				
Pedagogy	0.54***	-0.27***	0.41***	1.00			
Formalization	-0.14*	0.70***	-0.16**	-0.23***	1.00		
Progression	0.02	0.73***	0.04	-0.31***	0.55***	1.00	
Grade	0.18**	-0.45***	0.18**	0.30***	-0.40***	-0.30***	1.00

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). Correlations are calculated on teacher-subject-year means.

Lecture: The correlation between mean evaluation of teacher's pedagogy and mean evaluation of interest for the subject is 54 %. The correlation between mean evaluation of teacher's pedagogy and mean exam grade is 30 %.

Table 3.9 – Variance decomposition of evaluations and grades

	Inter-student variance	Inter-course variance
Pedagogy	86 %	14 %
Interest	92 %	8 %
Difficulty	74 %	26 %
Assiduousness	91 %	9 %
Formalization	91 %	9 %
Progression	89 %	11 %
Grades	84 %	16 %

Lecture: 86% of the total variance of evaluations of teacher's pedagogy is due to inter-student variance (intra-course variance). 14% is due to inter-course variance.

Table 3.10 – Estimation of the effect of mean grade on teacher’s pedagogy mean evaluation - IV with a change of teacher

Teacher’s pedagogy	mean			%very good			%very bad		
	OLS	FE	IV	OLS	FE	IV	OLS	FE	IV
Mean exam grade	0.077*** (0.022)	0.072* (0.037)	0.390 (0.315)	0.124 (0.236)	0.447 (0.326)	3.848 (2.822)	-0.762** (0.314)	-0.552 (0.424)	-8.038* (4.642)
Mean cc without exam	0.059*** (0.018)	0.094** (0.036)	0.323 (0.241)	0.093 (0.204)	0.293 (0.339)	2.806 (2.129)	-0.767*** (0.286)	-0.951** (0.438)	-6.571* (3.488)
Mean cc with exam	-0.001 (0.008)	0.005 (0.012)	0.056 (0.073)	0.219** (0.094)	-0.075 (0.161)	0.796 (0.670)	-0.086 (0.085)	0.157 (0.115)	-1.509 (1.165)
Intercept	1.803*** (0.276)	1.856*** (0.464)	-3.059 (4.496)	15.350*** (3.551)	15.749*** (3.896)	-32.441 (40.373)	23.670*** (4.229)	18.406*** (5.148)	128.997* (67.035)
Course characteristics	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
Teacher characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
After 2008 dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.11	0.77	.	0.20	0.73	.	0.12	0.72	.
Nb obs	485	485	59	485	485	59	485	485	59
Nb clusters	97	97	42	97	97	42	97	97	42

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). The unit of observation is a triplet teacher-course-year. Observations are weighted by the number of students in the unit. Standard errors in brackets are clustered at the course level. Course characteristics = dummies for type of subject, dummies for year 1, 2 or 3 of the school program, dummy for direct admission catch-up course. Teacher characteristics = dummy for male, dummy for first year of teaching. The FE specification contains teacher-course fixed effects. In the IV specification, instrumental variables = lagged mean exam grade, lagged mean continuous assessment grade without exam, lagged mean continuous assessment grade with exam, and proportion of students retaking the exam, in the case when the teachers changes between $t - 1$ and t .

Lecture: On average in the sample, 23.670% of students in a class rate the teacher’s pedagogy as very bad. A one point increase in mean exam grade is associated with a 0.762 decrease in this proportion. This coefficient is significant at the 5% significance level.

Table 3.11 – First stages with a change of teacher

First stage	Mean exam	Mean cc without exam	Mean cc with exam
Lag mean exam	0.258 (0.298)	-0.022 (0.293)	-0.447 (0.466)
Lag mean cc without exam	-0.214 (0.284)	0.495 (0.321)	-0.333 (0.329)
Lag mean cc with exam	-0.073 (0.119)	-0.022 (0.125)	0.574*** (0.175)
% retake	0.145* (0.085)	-0.249** (0.100)	0.135 (0.080)
Intercept	6.214 (4.868)	5.778 (5.466)	8.197 (6.777)
Course characteristics	Yes	Yes	Yes
Teacher characteristics	Yes	Yes	Yes
After 2008 dummy	Yes	Yes	Yes
R2	0.65	0.72	0.74
Fstat 1st stage	10	15	13
Nb obs	59	59	59
Nb clusters	42	42	42

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). The unit of observation is a triplet teacher-course-year. Observations are weighted by the number of students in the unit. Standard errors in brackets are clustered at the course level. Course characteristics = dummies for type of subject, dummies for year 1, 2 or 3 of the school program, dummy for direct admission course. Teacher characteristics = dummy for male, dummy for first year of teaching.

Lecture: A one point increase in mean exam grade the preceding year is associated with a 0.258 increase in current mean exam grade. This coefficient is not significantly different from zero.

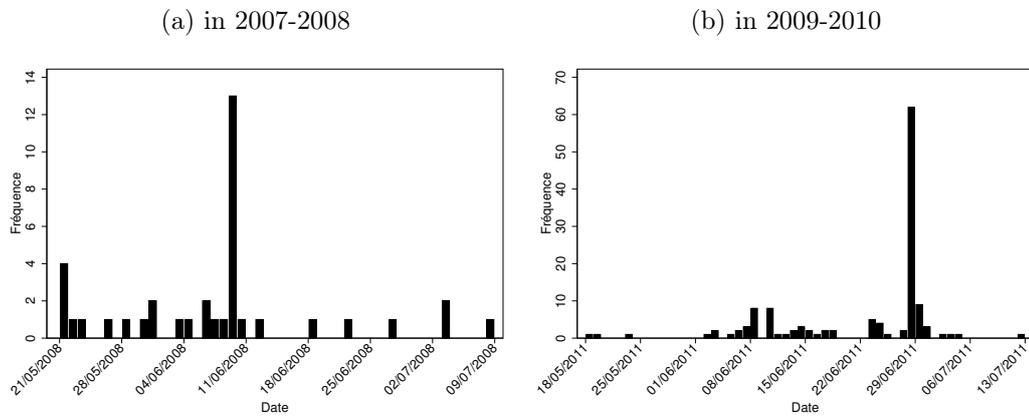
Table 3.12 – Estimation of the effect of mean grade on teaching assistant’s pedagogy mean evaluation

TA’s pedagogy	mean			%very good			%very bad		
	OLS	FE	IV	OLS	FE	IV	OLS	FE	IV
Mean exam grade	-0.005 (0.018)	0.007 (0.042)	-0.101* (0.057)	-0.287 (0.192)	0.050 (0.578)	-1.920*** (0.464)	-0.216 (0.198)	-0.037 (0.488)	0.390 (0.813)
Mean cc with exam	0.034* (0.018)	0.003 (0.021)	0.075 (0.060)	0.359 (0.289)	0.184 (0.408)	-0.145 (0.589)	-0.532 (0.344)	-0.453 (0.387)	1.258 (1.294)
Mean cc without exam	0.025 (0.022)	-0.144*** (0.044)	-0.006 (0.074)	-0.157 (0.296)	2.908*** (0.523)	-1.785*** (0.519)	-0.454 (0.369)	-1.942 (1.756)	1.703 (1.244)
Intercept	2.689*** (0.334)	3.115*** (0.536)	2.901** (1.141)	19.745*** (4.805)	14.052 (9.430)	43.365*** (8.848)	23.766*** (5.993)	21.935** (9.657)	-9.258 (20.637)
Course characteristics	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes
First year teaching	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
After 2008 dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.07	0.87	0.01	0.11	0.81	.	0.08	0.80	.
Nb obs	517	517	184	517	517	184	517	517	184
Nb clusters	28	28	25	28	28	25	28	28	25

Note: * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$). The unit of observation is a triplet teaching assistant-course-year. Observations are weighted by the number of students in the unit. Standard errors in brackets are clustered at the course level. Course characteristics = dummies for type of subject, dummies for year 1, 2 or 3 of the school program, dummy for direct admission catch-up course. The FE specification contains teaching assistant-course fixed effects. In the IV specification, instrumental variables = lagged mean exam grade, lagged mean continuous assessment grade without exam, lagged mean continuous assessment grade with exam, and proportion of students retaking the exam.

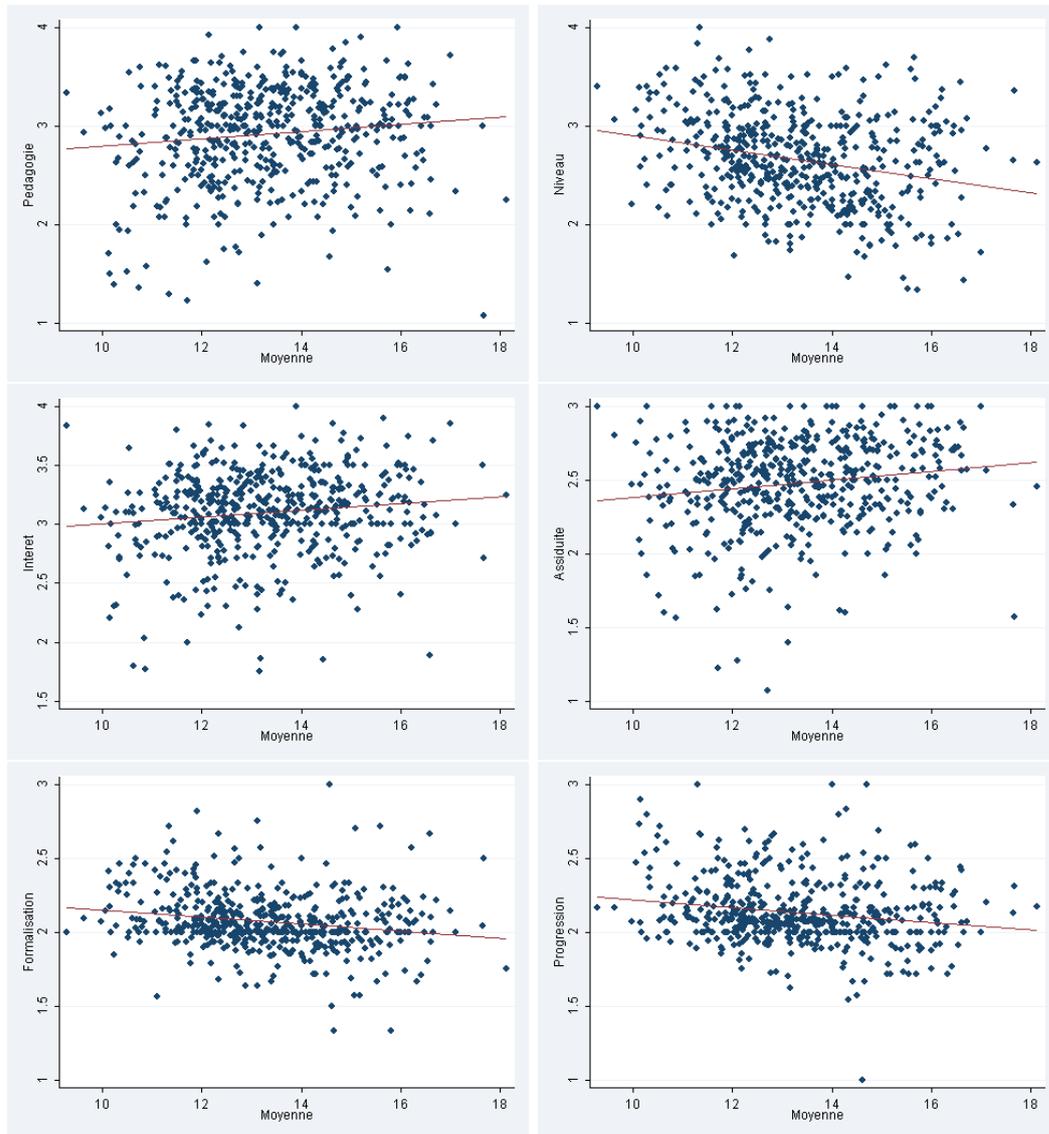
Lecture: On average in the sample, 23.766% of students in a tutorial class rate the teaching assistant’s pedagogy as very bad. A one point increase in mean exam grade is associated with a 0.216 decrease in this proportion. This coefficient is not significantly different from zero.

Figure 3.2 – Dates of evaluations



Lecture: These graphs present the histograms of the number of evaluations per day for one course in 2007-2008 (when evaluation was optional) and in 2009-2010 (when evaluation was compulsory).

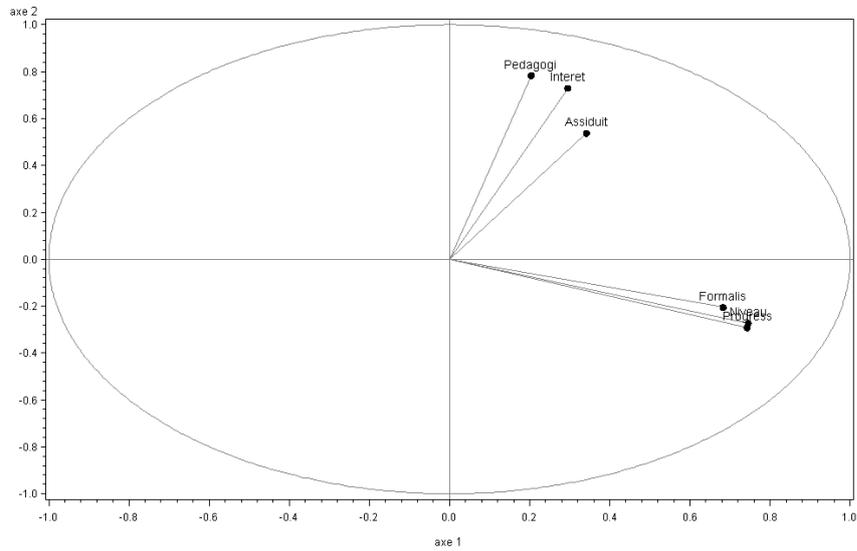
Figure 3.3 – Linear relationship between evaluations and exam grade



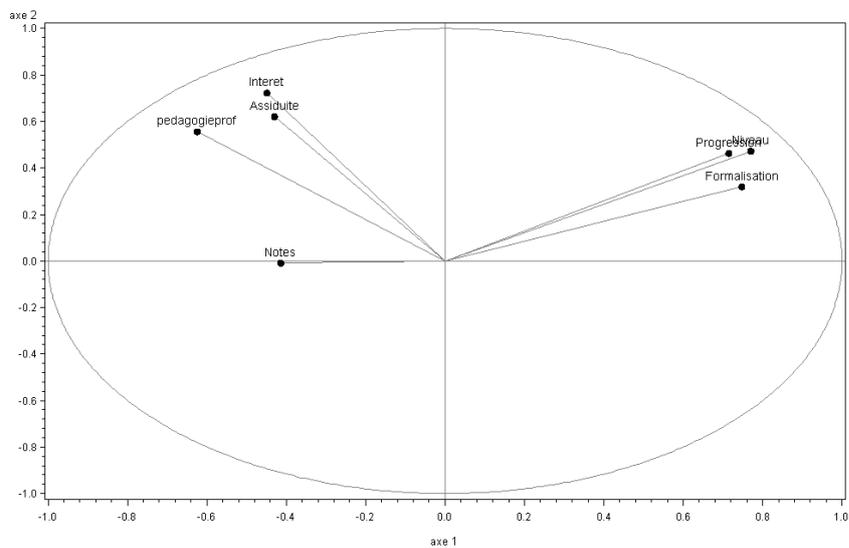
Note: These graphs plot by teacher-subject-year mean evaluation over mean exam grade. The lines represent the linear regression fit of mean evaluations over mean grade.

Figure 3.4 – Multivariate relationship between evaluations and exam grade

(a) At the individual level



(b) At the teacher-subject-year aggregate level



Note: These graphs represent the scatter plot of a principal component analysis, at the individual level (top), and at the teacher-subject-year aggregate level (bottom).

Conclusion

L'objectif principal de cette thèse était de tenir compte des réactions individuelles dans l'évaluation des dispositifs éducatifs. À travers trois analyses, nous avons étudié comment les individus ajustent leurs comportements en matière d'éducation en réponse à une modification des ressources scolaires ou à une modification de l'information relative à ces ressources. Nous avons proposé différentes méthodes économétriques permettant d'identifier séparément les stratégies individuelles et l'effet des ressources scolaires. Des données très riches, observées au niveau des individus, nous ont permis d'analyser de façon détaillée leurs comportements de scolarisation et d'évaluation. L'observation de la répartition géographique des établissements par rapport aux individus nous a par ailleurs permis d'étudier l'allocation spatiale des ressources scolaires.

Le premier chapitre a montré que les individus ajustent leur comportement de scolarisation aux dispositifs d'éducation prioritaire en changeant d'établissement. Vivre à proximité d'un établissement réseaux ambition réussite augmente d'environ 30 points de pourcentage la probabilité d'éviter cet établissement en se scolarisant dans un autre établissement public ou privé, pour les élèves traités de façon exogène. Ces ajustements sont observés dès la première année de mise en place du dispositif. Les stratégies de contournement des établissements concernés ne sont pas homogènes dans la population ; les individus les plus avantagés socialement et ceux qui ont une meilleure connaissance du système scolaire ont tendance à éviter les collèges labellisés RAR plus que les autres. Ils semblent répondre à un signal négatif sur la qualité des établissements. Lorsque l'on tient compte de l'évitement scolaire, on ne trouve globalement pas d'effet de la scolarisation en RAR sur les résultats des élèves à la fin du collège. Des effets différents sont observés pour les filles par rapport aux garçons et pour les élèves issus de catégories sociales plus avantagés par rapport aux autres. Ces effets sont cependant peu robustes au choix de l'échantillon d'étude. Cet

absence relative d'effet est vraisemblablement la résultante d'effets négatifs dûs à la fuite des élèves les plus favorisés socialement et d'effets bénéfiques du dispositif, qui, en moyenne, s'annulent.

Dans le deuxième chapitre, nous avons montré que les trajectoires scolaires des élèves sont modifiées par l'apparition de nouveaux lycées dans le voisinage. L'ouverture d'un nouveau lycée augmente d'environ 2 points de pourcentage la proportion d'élèves qui continuent en second cycle du secondaire. Cette augmentation est due à une plus grande proportion d'élèves qui continuent vers une seconde professionnelle. Les élèves qui sont contraints par l'offre scolaire locale sont donc ceux qui souhaiteraient poursuivre leur scolarité en voie professionnelle à la fin du collège. Les effets observés varient selon la taille de l'offre locale d'établissements. Ces résultats suggèrent que l'allocation spatiale des ressources est susceptible de créer des contraintes scolaires différentes sur le territoire. Ces contraintes semblent par ailleurs d'autant plus importantes pour les élèves dont le parcours scolaire est le plus fragile.

Le dernier chapitre a montré que les étudiants tiennent compte de la note donnée par l'enseignant lorsqu'ils évaluent ses qualités pédagogiques. Une augmentation d'un point de la note moyenne du cours à l'examen est associée à une diminution d'environ 1 point de pourcentage de la proportion d'élèves donnant de très mauvaises évaluations de la qualité pédagogique de l'enseignant. En outre, les étudiants utilisent l'information contemporaine dont ils disposent au moment où ils évaluent. Les étudiants donnent de moins bonnes évaluations après avoir passé l'examen et de meilleures évaluations après avoir obtenu leur note. Nos résultats suggèrent que les institutions d'enseignement doivent être attentives au moment auquel les cours sont évalués. Nous n'en concluons pas pour autant que les évaluations des enseignements par les étudiants ne devraient pas être utilisées par les institutions pour évaluer la formation. Les étudiants sont les mieux à même d'évaluer les qualités des cours qu'ils ont suivis.

Il est utile de noter que la quasi-totalité des résultats présentés dans cette thèse sont estimés uniquement pour les individus affectés par la modification des ressources. Il s'agit des « *compliers* », c'est-à-dire des individus dont le comportement change en réponse à un choc dans les ressources scolaires. Ainsi, dans le premier chapitre, les résultats sont valables uniquement pour les élèves qui sont scolarisés en RAR, parce que leur collège public le plus proche est juste au-dessus des seuils d'éligibilité et pour les élèves qui ne sont pas scolarisés en RAR, parce que leur collège public le plus proche est juste au-dessous des seuils. La méthode d'estimation utilisée ne nous

permet pas de tirer de conclusion pour les élèves qui ne sont pas dans l'une de ces deux situations, ni pour ceux dont l'établissement le plus proche est loin des seuils. Nous ne pouvons rien dire concernant les élèves qui auraient, de toute façon, contourné les établissements RAR, quelle qu'ait été la situation du collège le plus proche vis-à-vis des seuils d'éligibilité. De la même manière, dans le deuxième chapitre, nos résultats ne se généralisent pas aux élèves qui auraient de toute façon poursuivi leurs études en second cycle en l'absence de l'ouverture d'un nouveau lycée. Une des limitations des méthodes proposées ici est donc le caractère local des résultats. En contrepartie, l'avantage de ces méthodes est de mettre en évidence des effets causaux pour les élèves concernés, c'est-à-dire d'identifier les effets propres des dispositifs éducatifs, en tenant compte de leur endogénéité. Le caractère local des résultats présentés ici ne doit toutefois pas en limiter la portée ; ils permettent d'apporter des éclairages sur les réactions des individus à une modification des ressources scolaires et sur la façon dont il faut en tenir compte pour évaluer les dispositifs éducatifs.

Une autre limite de nos résultats est qu'ils ne permettent pas d'éliciter complètement les mécanismes par lesquels les individus réagissent à une modification des ressources scolaires. Les modèles dits en « forme réduite » utilisés ici n'explicitent pas quelles séquences de décisions conduisent par exemple les élèves à poursuivre en second cycle suite à l'ouverture d'un nouveau lycée, ni quelles seraient leurs préférences en termes d'orientation. Autrement dit, nous n'avons pas modélisé ici la structure des préférences des individus. Cependant, l'observation des modifications dans les comportements individuels permet de comprendre, de manière indirecte, les préférences individuelles en matière d'éducation. Le fait que les familles changent d'établissement en réponse au dispositif RAR nous apprend par exemple que les préférences en matière de choix d'établissement scolaire dépendent de l'information sur la qualité du collège le plus proche. De plus, l'effet de la révélation d'information dépend du milieu social. Ce travail représente ainsi une première étape dans la mise en place de modèles plus structurels. Autrement dit, avant de chercher à caractériser les préférences, il nous semble important d'essayer de les comprendre et de formuler des hypothèses testables.

Pour conclure, nous avons montré, dans cette thèse, que l'évaluation de l'impact des ressources scolaires dans la fonction de production de l'éducation nécessite de tenir compte des réactions individuelles. Les résultats présentés ici ne doivent pas dissuader les décideurs publics de mettre en place des politiques éducatives, mais les ajustements individuels ne peuvent pas être ignorés car ils ne sont pas aléatoires dans

la population. En outre, la dimension spatiale de l'allocation des ressources scolaires doit retenir l'attention des décideurs publics. D'une part, la localisation et la re-localisation des familles sont susceptibles de modifier l'effet attendu des politiques éducatives. D'autre part, les individus ne sont pas égaux devant l'allocation spatiale des ressources.

Bibliographie

- ANGRIST, J. D. et V. LAVY (1999). “Using Maimonides’ Rule to Estimate the Effect of Class Size on Scholastic Achievement”. *The Quarterly Journal of Economics* 114.2, p. 533–575.
- ATTANASIO, O. P., C. MEGHIR et A. SANTIAGO (2012). “Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA”. *The Review of Economic Studies* 79.1, p. 37–66.
- AVVISATI, F. et al. (2013). “Getting Parents Involved: A Field Experiment in Deprived Schools”. *The Review of Economic Studies*.
- BARROW, L. (2002). “School choice through relocation: evidence from the Washington, D.C. area”. *Journal of Public Economics* 86.2, p. 155–189.
- BAYER, P., F. FERREIRA et R. MCMILLAN (2007). “A Unified Framework for Measuring Preferences for Schools and Neighborhoods”. *Journal of Political Economy* 115.4, p. 588–638.
- BECKER, G. S. (1962). “Investment in Human Capital: A Theoretical Analysis”. *Journal of Political Economy* 70.5, p. 9–49.
- (1964). *Human Capital: A Theoretical Analysis with Special Reference to Education*. National Bureau for Economic Research, Columbia University Press, New York et London.
- BECKER, W. E., W. BOSSHARDT et M. WATTS (2012). “Revisiting How Departments of Economics Evaluate Teaching”. Working paper presented at the annual meetings of the American Economic Association.
- BECKER, W. E. et M. WATTS (1999). “How Departments of Economics Evaluate Teaching”. *American Economic Review* 89.2, p. 344–349.
- BEFFY, M. et L. DAVEZIES (2013). “Has the "Ambition Success" educational program achieved its ambition?” *Annales d’Économie et de Statistique* 111/112, p. 271–294.

- BELECHE, T., D. FAIRRISS et M. MARKS (2012). "Do course evaluations truly reflect student learning? Evidence from an objectively graded post-test". *Economics of Education Review* 31.5, p. 709–719.
- BERTRAND, M., E. DUFLO et S. MULLAINATHAN (2004). "How Much Should We Trust Differences-In-Differences Estimates?" *The Quarterly Journal of Economics* 119.1, p. 249–275.
- BLACK, S. E. (1999). "Do Better Schools Matter? Parental Valuation of Elementary Education". *The Quarterly Journal of Economics*, p. 577–599.
- BLACK, S. E. et S. MACHIN (2011). "Chapter 10 - Housing Valuations of School Performance". Sous la dir. de S. M. ERIC A. HANUSHEK et L. WOESSMANN. T. 3. *Handbook of the Economics of Education*. Elsevier, p. 485–519.
- BÉNABOU, R., F. KRAMARZ et C. PROST (2004). "Zones d'éducation prioritaire : quels moyens pour quels résultats? Une évaluation sur la période 1982-1992". *Économie et statistique* 380, p. 3–29.
- (2009). "The French zones d'éducation prioritaire: Much ado about nothing?" *Economics of Education Review* 28.3, p. 345–356.
- BORING, A. (2015). *Gender Biases in Student Evaluations of Teachers*. Working Paper. Observatoire Français des Conjonctures Économiques (OFCE).
- BORMAN, G. D. et J. V. D'AGOSTINO (1996). "Title I and Student Achievement: A Meta-Analysis of Federal Evaluation Results". *Educational Evaluation and Policy Analysis* 18.4, p. 309–326.
- BRAGA, M., M. PACCAGNELLA et M. PELLIZZARI (2011). "Evaluating students' evaluations of professors". *IZA Discussion Paper No.* 5620.
- (2014). "Evaluating students' evaluations of professors". *Economics of Education Review* 41, p. 71–88.
- BRESSOUX, P., F. KRAMARZ et C. PROST (2009). "Teachers' Training, Class Size and Students' Outcomes: Learning from Administrative Forecasting Mistakes". *The Economic Journal* 119.536, p. 540–561.
- BURDE, D. et L. LINDEN (2013). "Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools". *American Economic Journal: Applied Economics* 5.3, p. 27–40.
- CAILLE, J.-P. (2001). "Les collégiens de ZEP à la fin des années quatre-vingt-dix : caractéristiques des élèves et impact de la scolarisation en ZEP sur la réussite". *Education et formations* 61, p. 111–140.

-
- CAILLE, J.-P., L. DAVEZIES et M. GARROUSTE (2015). “Les réseaux ambition réussite. Une analyse par régression sur discontinuité”. *Revue économique*, forthcoming.
- CARRELL, S. E. et J. E. WEST (2010). “Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors”. *Journal of Political Economy* 118.3, p. 409–432.
- CHUMACERO, R. A., D. GÓMEZ et R. D. PAREDES (2011). “I would walk 500 miles (if it paid): Vouchers and school choice in Chile”. *Economics of Education Review* 30.5, p. 1103–1114.
- COLEMAN, J. S. et al. (1966). *Equality of Educational Opportunity*. Rapp. tech. U.S. Department of Health, Education, et Welfare, National Center For Educational Statistics, p. 1066–5684.
- CUNHA, F. et J. HECKMAN (2007). “The Technology of Skill Formation”. *American Economic Review* 97.2, p. 31–47.
- D’APOLLONIA, S. et P. C. ABRAMI (1997). “Navigating Student Ratings of Instruction”. *American psychologist* 52.11, p. 1198–1208.
- DAS, J. et al. (2013). “School Inputs, Household Substitution, and Test Scores”. *American Economic Journal: Applied Economics* 5.2, p. 29–57.
- DAVEZIES, L. et M. GARROUSTE (2014). *More harm than good? Sorting effects in a compensatory education program*. Working Paper 2014-42. CREST Working Paper.
- DAVEZIES, L. et T. LE BARBANCHON (2014). “Regression Discontinuity Design with Continuous Measurement Error in the Running Variable”. 2014-27.
- DEJEAN, J. (2002). *L’évaluation de l’enseignement dans les universités françaises*. Rapp. tech. Haut Conseil de l’évaluation de l’école.
- DEMING, D. J. et al. (2014). “School Choice, School Quality, and Postsecondary Attainment”. *American Economic Review* 104.3, p. 991–1013.
- DEWEY, J., T. A. HUSTED et L. W. KENNY (2000). “The ineffectiveness of school inputs: a product of misspecification?” *Economics of Education Review* 19.1, p. 27–45.
- DICKERSON, A. et S. MCINTOSH (2013). “The Impact of Distance to Nearest Education Institution on the Post-compulsory Education Participation Decision”. *Urban Studies* 50.4, p. 742–758.

- DOBBIE, W. et R. G. FRYER (2011). “Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children’s Zone”. *American Economic Journal: Applied Economics* 3.3, p. 158–187.
- DUFLO, E. (2001). “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment”. *American Economic Review* 91.4, p. 795–813.
- EPPLE, D., T. ROMER et H. SIEG (2001). “Interjurisdictional Sorting and Majority Rule: An Empirical Analysis”. *Econometrica* 69.6, p. 1437–1465.
- EPPLE, D. N. et R. ROMANO (2003). “Neighborhood Schools, Choice, and the Distribution of Educational Benefits”. *The Economics of School Choice*. Sous la dir. de C. M. HOXBY. University of Chicago Press, p. 227–286.
- EWING, A. M. (2012). “Estimating the impact of relative expected grade on student evaluations of teachers”. *Economics of Education Review* 31.1, p. 141–154.
- FACK, G. et J. GRENET (2010). “When do better schools raise housing prices? Evidence from Paris public and private schools”. *Journal of Public Economics* 94.1, p. 59–77.
- (2012). *Rapport d’évaluation de l’assouplissement de la carte scolaire*. Rapp. tech. Centre pour la recherche économique et ses applications.
- (2013). “Les effets de l’assouplissement de la carte scolaire dans l’éducation prioritaire”. *Éducation et Formations* 83, p. 25–37.
- FALCH, T., P. LUJALA et B. STRØM (2013). “Geographical constraints and educational attainment”. *Regional Science and Urban Economics* 43.1, p. 164–176.
- FERREYRA, M. M. (2007). “Estimating the Effects of Private School Vouchers in Multidistrict Economies”. *American Economic Review* 97.3, p. 789–817.
- FIGLIO, D. N. et M. E. LUCAS (2004). “What’s in a Grade? School Report Cards and the Housing Market”. *American Economic Review* 94.3, p. 591–604.
- FILMER, D. (2007). “If you build it, will they come? School availability and school enrolment in 21 poor countries”. *Journal of Development Studies* 43.5, p. 901–928.
- FRANZ, W.-J. I. (2010). “Grade inflation under the threat of students’ nuisance: Theory and evidence”. *Economics of Education Review* 29.3, p. 411–422.
- FRENETTE, M. (2009). “Do universities benefit local youth? Evidence from the creation of new universities”. *Economics of Education Review* 28.3, p. 318–328.
- FRIESEN, J. et al. (2012). “How do school ‘report cards’ affect school choice decisions?” *Canadian Journal of Economics/Revue canadienne d’économique* 45.2, p. 784–807.

-
- GIBBONS, S. et A. VIGNOLES (2012). “Geography, choice and participation in higher education in England”. *Regional Science and Urban Economics* 42.1-2, p. 98–113.
- GOUX, D., M. GURGAND et E. MAURIN (2015). “Adjusting Your Dreams? Highschool Plans and Dropout Behaviour”. *Economic Journal*, forthcoming.
- GOUX, D. et E. MAURIN (2007). “Close Neighbours Matter: Neighbourhood Effects on Early Performance at School”. *The Economic Journal* 117.523, p. 1193–1215.
- GREENWALD, A. G. (1997). “Validity Concerns and Usefulness of Student Ratings of Instruction”. *American Psychologist* 52.11, p. 1182.
- GREENWALD, A. G. et G. M. GILLMORE (1997). “Grading Leniency is a Removable Contaminant of Student Ratings”. *American psychologist* 52.11, p. 1209.
- GREENWALD, R., L. V. HEDGES et R. D. LAINE (1996). “The Effect of School Resources on Student Achievement”. *Review of Educational Research* 66.3, p. 361–396.
- GRIFFITH, A. et D. ROTHSTEIN (2009). “Can’t get there from here: The decision to apply to a selective college”. *Economics of Education Review* 28.5, p. 620–628.
- HAHN, J., P. TODD et W. Van der KLAUW (2001). “Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design”. *Econometrica* 69.1, p. 201–09.
- HANDA, S. (2002). “Raising primary school enrolment in developing countries: The relative importance of supply and demand”. *Journal of Development Economics* 69.1, p. 103–128.
- HANSEN, C. B. (2007). “Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects”. *Journal of Econometrics* 140.2, p. 670–694.
- HANUSHEK, E. A. (1981). “Throwing Money at Schools”. *Journal of Policy Analysis and Management* 1.1, p. 19–41.
- (1986). “The Economics of Schooling: Production and Efficiency in Public Schools”. *Journal of Economic Literature* 24.3, p. 1141–1177.
- (1989). “The Impact of Differential Expenditures on School Performance”. *Educational Researcher* 18.4, p. 45–51+62.
- (1997). “Assessing the Effects of School Resources on Student Performance: An Update”. *Educational evaluation and policy analysis* 19.2, p. 141–164.
- (1998). “Conclusions and Controversies about the Effectiveness of School Resources”. *Economic Policy Review* 4.1.

- HANUSHEK, E. A. (2003). "The Failure of Input-based Schooling Policies". *The Economic Journal* 113.485, F64–F98.
- HASTINGS, J., T. KANE et D. STAIGER (2009). "Heterogeneous Preferences and the Efficacy of Public School Choice". *unpublished working paper*.
- HASTINGS, J. S. et J. M. WEINSTEIN (2008). "Information, School Choice, and Academic Achievement: Evidence from Two Experiments". *The Quarterly Journal of Economics* 123.4, p. 1373–1414.
- HEDGES, L. V., R. D. LAINE et R. GREENWALD (1994). "Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes". *Educational Researcher* 23.3, pp. 5–14.
- HOFFMANN, F. et P. OREOPOULOS (2009). "Professor Qualities and Student Achievement". *The Review of Economics and Statistics* 91.1, p. 83–92.
- HOROWITZ, J. L. et C. F. MANSKI (1995). "Identification and Robustness with Contaminated and Corrupted Data". *Econometrica: Journal of the Econometric Society* 63.2, p. 281–302.
- HOUTENVILLE, A. J. et K. S. CONWAY (2008). "Parental Effort, School Resources, and Student Achievement". *Journal of Human Resources* 43.2, p. 437–453.
- HOBY, C. M. (2000). "The Effects of Class Size on Student Achievement: New Evidence from Population Variation". *The Quarterly Journal of Economics* 115.4, p. 1239–1285.
- HSIEH, C.-T. et M. URQUIOLA (2006). "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program". *Journal of Public Economics* 90.8–9, p. 1477–1503.
- IMBENS, G. W. et J. D. ANGRIST (1994). "Identification and Estimation of Local Average Treatment Effects". English. *Econometrica* 62.2, p. 467–475.
- IMBENS, G. W. et T. LEMIEUX (2008). "Regression discontinuity designs: A guide to practice". *Journal of Econometrics* 142.2, p. 615–635.
- IMBENS, G. et K. KALYANARAMAN (2012). "Optimal Bandwidth Choice for the Regression Discontinuity Estimator". *The Review of Economic Studies* 79.3, p. 933–959.
- ISELY, P. et H. SINGH (2005). "Do Higher Grades Lead to Favorable Student Evaluations?" *The Journal of Economic Education* 36.1, p. 29–42.
- KLAAUW, W. van der (2008). "Breaking the link between poverty and low student achievement: An evaluation of Title I". *Journal of Econometrics* 142.2, p. 731–756.

-
- KLITGAARD, R. (1990). *Controlling Corruption*. University of California Press.
- KRAUTMANN, A. C. et W. SANDER (1999). “Grades and student evaluations of teachers”. *Economics of Education Review* 18.1, p. 59–63.
- KRUEGER, A. B. (1999). “Experimental Estimates of Education Production Functions”. *The Quarterly Journal of Economics* 114.2, p. 497–532.
- KRUEGER, A. B. (2003). “Economic Considerations and Class Size”. *The Economic Journal* 113.485, F34–F63.
- LADD, H. F. (1994). “Spatially Targeted Economic Development Strategies: Do They Work?” *Cityscape: A Journal of Policy Development and Research* 1.1, p. 193–218.
- LAVY, V. (2010). “Effects of Free Choice Among Public Schools”. *The Review of Economic Studies* 77.3, p. 1164–1191.
- (2015). *Long Run Effects of Free School Choice: College Attainment, Employment, Earnings, and Social Outcomes at Adulthood*. Working Paper 20843. National Bureau of Economic Research.
- MACHIN, S., S. McNALLY et C. MEGHIR (2004). “Improving Pupil Performance in English Secondary Schools: Excellence in Cities”. *Journal of the European Economic Association* 2.2-3, p. 396–405.
- (2010). “Resources and Standards in Urban Schools”. *Journal of Human Capital* 4.4, p. 365–393.
- MARSH, H. W. et L. A. ROCHE (1997). “Making Students’ Evaluations of Teaching Effectiveness Effective: The Critical Issues of Validity, Bias, and Utility”. *American Psychologist* 52.11, p. 1187–1197.
- MCCRARY, J. (2008). “Manipulation of the running variable in the regression discontinuity design: A density test”. *Journal of Econometrics* 142.2, p. 698–714.
- MEURET, D. (1994). “L’efficacité de la politique des zones d’éducation prioritaire dans les collèges”. fr. *Revue française de pédagogie* 109.1, p. 41–64.
- MIDDLETON, S. et al. (2005). *Evaluation of Education Allowance Pilots: young people aged 16 to 19 years*. Rapp. tech. Research report, Department for education et skills.
- MINCER, J. A. (1974). *Schooling, Experience, and Earnings*. National Bureau of Economic Research.
- MOISAN, C. et J. SIMON (1997). *Les Déterminants de la réussite scolaire en zone d’éducation prioritaire*. Rapp. tech. Inspection générale de l’administration de

- l'Éducation nationale. Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche.
- MORETTI, E. (2011). "Social Learning and Peer Effects in Consumption: Evidence from Movie Sales". *Review of Economic Studies* 78.1, p. 356–393.
- MURAT, F. (1998). "Les différentes façons d'évaluer le niveau des élèves en fin de collège". *Éducation et formations* 53, p. 35–49.
- MURAT, F. et M. THAUREL-RICHARD (2013). "Évolution des caractéristiques des collèges durant la mise en oeuvre de l'assouplissement de la carte scolaire de 2007". *Éducation et Formations* 83, p. 11–20.
- NECHYBA, T. (2003). "Public School Finance and Urban School Policy: General Versus Partial Equilibrium Analysis". *Brookings-Wharton Papers on Urban Affairs*, p. 139–170.
- NEUMARK, D. et H. SIMPSON (2015). "Place-Based Policies". *Handbook of Regional and Urban Economics*. Sous la dir. de J. V. H. GILLES DURANTON et W. C. STRANGE. T. 5. Handbook of Regional and Urban Economics. Elsevier, p. 1197–1287.
- OECD (2013). *PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices*. Rapp. tech. OECD.
- PEI, Z. (2011). "Regression Discontinuity Design with Measurement Error in the Assignment Variable". *Unpublished manuscript, Department of Economics, University of Princeton*.
- PIKETTY, T. et M. VALDENAIRE (2006). "L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français: estimations à partir du panel primaire 1997 et du panel secondaire 1995". *Les dossiers évaluations et statistiques* 173.
- POP-ELECHES, C. et M. URQUIOLA (2013). "Going to a Better School: Effects and Behavioral Responses". *American Economic Review* 103.4, p. 1289–1324.
- RUBIN, D. B. (1974). "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies". *Journal of Educational Psychology* 66.5, p. 688–701.
- SCHULTZ, T. P. (2004). "School subsidies for the poor: evaluating the Mexican Progresa poverty program". *Journal of Development Economics* 74.1. New Research on Education in Developing Economies, p. 199–250.
- SCHULTZ, T. W. (1961). "Investment in Human Capital". *The American Economic Review* 51.1, pp. 1–17.

-
- SHAPIRO, J. et J. MORENO TREVIÑO (2004). “Compensatory Education for Disadvantaged Mexican Students: An Impact Evaluation Using Propensity Score Matching”. *World Bank Policy Research Working Paper* 3334.
- SPIESS, C. et K. WROHLICH (2010). “Does distance determine who attends a university in Germany?” *Economics of Education Review* 29.3, p. 470–479.
- TIEBOUT, C. M. (1956). “A Pure Theory of Local Expenditures”. *Journal of Political Economy* 64.5, p. 416–424.
- TOKMAN, A. (2002). *Evaluation of the P900 Program: A Targeted Education Program for Underperforming Schools*. Central Bank of Chile Working Papers 170. Central Bank of Chile.
- URQUIOLA, M. (2005). “Does School Choice Lead to Sorting? Evidence from Tiebout Variation”. *American Economic Review* 95.4, p. 1310–1326.
- URQUIOLA, M. et E. VERHOOGEN (2009). “Class-Size Caps, Sorting, and the Regression-Discontinuity Design”. *American Economic Review* 99.1, p. 179–215.
- VIGNOLES, A. et al. (2000). *The Relationship between Resource Allocation and Pupil Attainment: A Review*. CEE Discussion Papers 0002. Centre for the Economics of Education, LSE.
- WEINBERG, B. A., M. HASHIMOTO et B. M. FLEISHER (2009). “Evaluating teaching in higher education”. *The Journal of Economic Education* 40.3, p. 227–261.
- WÖSSMANN, L. (2005). “Educational production in Europe”. *Economic Policy* 20.43, p. 446–504.
- WÖSSMANN, L. (2003). “Schooling Resources, Educational Institutions and Student Performance: the International Evidence”. *Oxford Bulletin of Economics and Statistics* 65.2, p. 117–170.
- YU, P. (2012). “Identification of Treatment Effects in Regression Discontinuity Designs with Measurement Error”.

Titre Ressources scolaires et réactions individuelles. Trois essais en économie de l'éducation.

Résumé L'objectif de cette thèse est d'étudier l'interdépendance des ressources scolaires et des ressources individuelles dans la production du capital humain. À travers trois cas d'études sur données françaises, différentes méthodes sont proposées pour analyser l'effet des politiques éducatives tout en tenant compte des comportements des individus. Le premier chapitre montre que les politiques d'éducation prioritaire sont susceptibles de conduire les familles, notamment les plus avantagées socialement, à contourner les établissements traités. Ces stratégies d'évitement sont de nature à contrebalancer les effets des moyens supplémentaires sur les résultats des élèves. Le deuxième chapitre cherche à savoir si les choix d'orientation des élèves sont contraints par l'offre scolaire locale. Nous montrons que l'ouverture d'un nouveau lycée augmente la proportion d'élèves qui poursuivent leurs études dans le second cycle, notamment en voie professionnelle. Enfin, le troisième chapitre montre que les étudiants tiennent compte de l'information contemporaine dont ils disposent sur leurs notes lorsqu'ils évaluent les qualités pédagogiques de leurs enseignants.

Mots-clés Ressources scolaires, Politiques éducatives, Éducation prioritaire, Choix d'établissement, Allocation spatiale, Évaluation des enseignements par les étudiants

Title School Resources and Individual Responses. Essays in Economics of Education

Abstract The main objective of this thesis is to study the interdependence of school resources and individual resources in human capital production. In three empirical analyses using French data, various methods are proposed to evaluate the impact of educational policies, taking individual behaviors into account. The first chapter shows that compensatory education policies may result in individual sorting. The fact that socially more advantaged families tend to avoid treated schools cancels out additional resources positive effects on academic achievement. The second chapter examines whether pupils are constrained by local school supply. We find that opening a new high school increases the proportion of pupils who continue in upper secondary education, particularly in a vocational track. Studying how students evaluate teaching, the third chapter shows that they take contemporaneous information about their grades into account when they evaluate teachers' pedagogical qualities.

Keywords School resources, Educational policies, Compensatory education, School choice, Spatial distribution, Student evaluation of teaching