



**HAL**  
open science

# Market for education and student achievement

Alexis Le Chapelain

► **To cite this version:**

Alexis Le Chapelain. Market for education and student achievement. Economics and Finance. Institut d'études politiques de paris - Sciences Po, 2014. English. NNT : 2014IEPP0040 . tel-03510563

**HAL Id: tel-03510563**

**<https://theses.hal.science/tel-03510563>**

Submitted on 4 Jan 2022

**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.

Institut d'Études Politiques de Paris  
ÉCOLE DOCTORALE DE SCIENCES PO  
Programme doctoral en économie  
Département d'économie de Sciences Po

Doctorat en Sciences économiques

Market for Education and Student Achievement

Alexis Le Chapelain

Thèse dirigée par Jean-Marc Robin, Professeur des Universités,  
Sciences Po Paris

Soutenue à Paris, le 12 décembre 2014

**Jury :**

Denis Fougère, Directeur de recherche, CNRS / GRECSTA

Robert Gary-Bobo, Professeur des Universités, Université Paris 1 Panthéon-Sorbonne (rapporteur)

Marc Gurgand, Directeur de recherche, CNRS /PSE (rapporteur)

Jean-Marc Robin, Professeur des Universités, Sciences Po Paris

Olivier Tercieux, Directeur de recherche, CNRS /PSE

Etienne Wasmer, Professeur des Universités, Sciences Po Paris



A ma mère



## Remerciements

Alors que cette thèse s'achève, il me faut maintenant remercier tous ceux qui, durant trois ans, m'ont permis de mener cette entreprise à bien. Je n'aurais pu achever ce travail sans tous ceux qui, de manière petite ou grande, consciente ou non, m'ont aidé. Je tiens tout d'abord à remercier mon directeur de thèse Jean-Marc Robin, qui m'a encadré d'abord en mémoire puis en thèse pendant près de quatre ans. Il a toujours été disponible pour répondre à mes questions, lire mes articles et projets de recherches, même les plus malavisés, et m'a toujours offert un soutien matériel sans faille, tout en me laissant une grande liberté intellectuelle et une grande liberté de recherche.

Je voudrais aussi remercier Robert Gary-Bobo et Marc Gurgand, qui ont bien voulu être les rapporteurs de cette thèse, ainsi que Denis Fougère, Olivier Tercieux et Etienne Wasmer, qui ont accepté d'être membre du jury. J'apprécie leurs efforts pour lire et commenter une thèse composée pour beaucoup de travaux encore inachevés. Leurs conseils et leurs commentaires me seront précieux pour les améliorer.

Je n'aurais pu rédiger cette thèse sans l'aide de nombreux professeurs, aussi bien ceux dont j'ai été l'élève pendant mes études que ceux qui m'ont conseillé et aidé durant mes années de thèse. Ils sont trop nombreux pour les nommer tous. Je voudrais cependant remercier plus particulièrement Jérôme Pouyet, qui encadra mon stage de recherche à la fin de ma première année de master en économie avec beaucoup de patience et de pédagogie. Je remercie également Yann Algan et Quoc-Anh Do, avec qui j'ai eu l'opportunité de collaborer sur des projets de recherche commun, et au contact de qui j'ai beaucoup appris. Parmi les professeurs du Département d'économie de Sciences Po, je remercie plus particulièrement Alfred Galichon, Sidartha Gordon, Guy Laroque, Thierry Mayer pour le temps qu'ils ont pris pour lire mes travaux, me recevoir et me conseiller. J'ai bénéficié également des conseils des chercheurs en visite à Sciences Po, notamment Antonio Cabralès, Victor Couture, Peter Arcidiaconno, Johannès Hörner, Yann Bramoullé et Antonio Gavazza, que je remercie ici.

Ecrire cette thèse m'a aussi donné l'opportunité de voyager, de fréquenter d'autres institutions, et de rencontrer des professeurs et doctorants qui m'ont beaucoup aidé. Je voudrais tout d'abord remercier l'ensemble du personnel de la Direction de l'Évaluation, de la Prospective et de la Performance (DEPP) de l'Éducation nationale, qui m'ont accueilli en stage et aidé à collecter les données utilisées pour les parties 2 et 3 de cette thèse. Je voudrais en particulier remercier Cédric Afsa, qui m'a autorisé à accéder aux données et suggéré le projet de recherche réalisé dans la partie 3, et Pierrette Briant, qui m'a donné toutes les informations pour comprendre et interpréter les données. Je remercie également Son-Thierry Ly de PSE, avec qui j'ai eu des discussions fructueuses, Xiaofei Ma du Crest, Léonard Moulin de Paris XIII, et Michal Fabinger de l'université de Tokyo, ainsi que Thomas Lindemann de Versailles-Saint-Quentin.

Au delà des professeurs du département d'économie, j'ai beaucoup profité du soutien

apporté par Sciences Po à ses doctorants et à la recherche. Je remercie donc chaleureusement Pilar Calvo, Cécile Markarian, Claudine Lamaze, Cathy Bénard, Leïla Brunoir, Alain Besoin, et tout l'équipe administrative du département d'économie et de l'école doctorale. Le chapitre 3 de cette thèse a pu être rédigé grâce à des données mis à ma disposition par la direction des études et de la scolarité dont je remercie le personnel avec gratitude, et plus particulièrement Myriam Dubois-Monkachi, Pascale Leclercq, Stéphane Auzanneau et Angélique Pawlowsky, qui ont pris beaucoup de leurs temps pour m'expliquer le fonctionnement du collège universitaire de Sciences Po, et pour créer et mettre à disposition les données. Je remercie également Richard Audoly, avec qui j'ai eu l'occasion de travailler sur ces données.

Durant mes années de thèse, j'ai eu l'occasion de cotoyer de nombreux doctorants qui m'ont permis de travailler dans un environnement chaleureux et agréable. Je remercie donc les anciens, Lisa, Elizaveta, François, Marion, Thomas, Gong, Camille, Marlène, Etienne L., Dilan, Gabriel, Ursula, Maxime. J'ai beaucoup bénéficié de l'interaction avec les membres de ma cohorte, et je remercie donc chaleureusement Camilo, Elisa, Jules, Meraj, Jamil, et Guillaume V., qui, je n'en doute pas, défendront bientôt leur thèse avec succès. Alors que j'ai commencé par être un jeune doctorant ayant plein de question à poser aux aînés, je me suis retrouvé sans m'en rendre compte être l'un des plus vieux du groupe... Je remercie donc tous les nouveaux (et moins nouveaux) doctorants qui ont su garder la même bonne humeur que leurs prédécesseurs, Zinedinne, Clément, Johannes, Guillaume C., Eve, Pierre C-C, Pierre D, Assia, Marta, Etienne F., Dylan, Pierre G., Florian, Jean-Louis, Charles-Louis, Emanuele, Andrea, Mathilde, Valeria, Joanne, Ivan, Léa, Corentin, Quentin, Paul, Jean-Baptiste, Guillaume W., Clara, Eiji et Ricardo, ainsi que Jorge, venu nous visiter depuis Columbia, et Christine, qui a accepté de m'aider sur certains projets de recherche. Je remercie aussi les ATER et les Postdoc qui sont venus compléter nos rangs ces dernières années, et en particulier Clément, Xavier, Morgane, Adrien, et Arthur. Je ne peux oublier mes collègues sociologues et politologues, et je remercie donc Frédéric et Adrien pour nos nombreuses conversations devant la machine à café.

Je remercie mes amis et mes proches, Sarah, Hye Yun, Raphaëlle, Diego, Quynh, Annabara, Thiên Anh, Tomoko, et ma soeur Camille, qui devrait bientôt devenir le second docteur de la famille.

Et avant tout, je remercie ma mère. Merci Maman





## Résumé

Cette thèse vise à discuter de la pertinence de l'introduction de mécanismes de marché au sein du système éducatif. En particulier, elle s'interroge sur la capacité de ces mécanismes à améliorer la réussite des élèves. En effet, partant du constat que les systèmes éducatifs publics traditionnels sont caractérisés par la présence d'un monopole des écoles sur leurs élèves, de nombreux chercheurs et hommes politiques ont plaidé pour une dérégulation du marché éducatif, susceptible de permettre l'entrée d'acteurs plus efficaces, de créer une pression compétitive sur les écoles, de leur donner plus d'autonomie, et de permettre un meilleur appariement entre élèves et écoles.

La première partie de cette thèse présente des travaux portant sur l'appariement entre établissement scolaires et élèves. Le chapitre 2 présente un modèle de formation de préférence pour les écoles et d'appariement en présence de préférences pour les pairs. Ils montrent qu'une telle interdépendance entre les préférences individuelles réduit l'efficacité des mécanismes d'appariement les plus courants en augmentant la corrélation des choix soumis par les élèves au planificateur. Une simple modification de ces mécanismes permet d'améliorer leur efficacité, tout en réduisant la ségrégation entre élèves. Le chapitre 3 présente une étude réalisée sur un programme de discrimination positive mis en œuvre par une grande école française. Il montre que ce programme a des effets mitigés dans la mesure où un nombre non négligeable de bénéficiaires du programme échoue, et que ceux qui réussissent à poursuivre leurs études ne parviennent que partiellement à combler l'écart académique qui les sépare de la moyenne des autres élèves. De plus, l'établissement examiné semble ne pas prendre en compte toute l'information disponible lors du recrutement des élèves, et admet donc des élèves dont la probabilité d'échec est élevée.

La seconde partie de cette thèse porte sur la mesure des effets de pairs. En effet, la présence de tels effets est cruciale dans le fonctionnement des marchés éducatifs. Le chapitre 5 estime les effets de pairs au sein des collèges français en utilisant la variation de la taille des classes. En effet, cela crée une faible variation dans les groupes de pairs auxquelles font face les élèves, et qui permet d'identifier l'impact que leurs camarades ont sur eux. En utilisant une base de données de très grande taille, cette méthode permet d'estimer des effets de pairs significatifs et de taille non négligeable.

La troisième partie de cette thèse s'intéresse à la rémunération et à l'efficacité des enseignants en France. Le chapitre 7 montre que la hiérarchie est partiellement capable d'identifier les meilleurs enseignants, mais que sur le long terme, ceux-ci ne sont pas mieux payés que les autres. Le chapitre 7 documente aussi une variation non négligeable d'efficacité entre enseignants, comparables à celle trouvée dans les études internationales, et montre que ces différences d'efficacité n'ont que peu de pouvoir prédictif à l'aide des variables observables telles que l'expérience ou bien le genre.

La dernière partie discute des effets pro-compétitifs des politiques de libéralisation scolaire, et plus particulièrement de celles faisant évoluer le système de sectorisation vers

des systèmes de libre choix. Le chapitre 9 présente ainsi un modèle de concurrence entre écoles qui met en évidence l'importance que revêt le rôle des pairs dans l'évaluation de la qualité des écoles dans la compétition entre les écoles. En effet, quand les familles accordent beaucoup d'importance aux pairs, et que le système scolaire est stratifié en école plus ou moins prestigieuse, les écoles regroupant les meilleurs élèves, au sommet de la hiérarchie, acquiert un pouvoir de marché qui réduit leurs incitations à fournir des efforts. En augmentant la différenciation verticale entre école, une politique de libéralisation du choix scolaire peut donc avoir pour effet de réduire la compétition entre les écoles, et de réduire le niveau d'effort exercé.

Les travaux originaux présentés dans cette thèse, ainsi que la discussion de la littérature existante sur le sujet qui y est effectuée permettent donc de conclure que les mécanismes de marché ne sont pas une solution efficace aux problèmes de performances et d'échec scolaire auxquels les systèmes éducatifs des pays développés doivent faire face. En effet, les spécificités du marché de l'éducation, où sont présents effets de pairs, information imparfaite, et rationalité limitée, rende les systèmes publics d'éducation aussi efficace, sinon plus, que les systèmes privés.

Cependant, comme le montre la troisième partie, les systèmes publics font face à des problèmes d'agence, et à des difficultés à gérer et récompenser leur personnel au mieux. Il est donc nécessaire de trouver des mécanismes alternatifs permettant d'améliorer l'efficacité de l'offre d'éducation.





## Summary

This dissertation aims at discussing the relevance of the introduction of market mechanisms in the education system. More specifically, it examines the ability of these mechanisms to improve student achievement. Indeed, starting from the fact that traditional public education system are associated with the presence of a monopoly of schools on their students, many researchers and politicians argued in favor of a deregulation of the education market, thinking that it would enable more efficient actors to enter, that it would create more competitive pressure on existing schools, that it would give schools more autonomy, and that it would lead to a better matching between students and schools.

The first part of this dissertation presents works discussing the matching between teachers and students. The second chapter proposes a model of preference formation for schools, and of matching with the existence of preference for peers. It shows that such interdependence across preferences reduces the efficiency of the most currently used matching mechanisms, by increasing the correlation of the choices submitted by students to the planner. A simple change in these mechanisms enables to improve their efficiency, while reducing segregation across students.

The third chapter displays a study on an affirmative action program implemented by a French “*grande écoles*”. It shows that this program has mixed effects, since a significant part of its beneficiaries fail, and that those who succeed and can pursue their study further can only incompletely bridge the academic gap with the other students. Moreover, the university implementing the program seems unable to fully take the available information into account, and thus admits students whose risk of failure is high.

The topic of the second part of this dissertation is the measure of peer effects. Indeed, the presence of such peer effects is fundamental in the functioning of educational markets. The fifth chapter estimates peer effects in French junior high schools by using class size variation. Indeed, class size creates a small variation in the peer groups that students have to face, which enables to identify the impact that their classmates have on them. By using a very large dataset, this method enables to estimate significant and non trivial peer effects.

The third part of the dissertations focuses on the compensation and the efficiency of teachers in France. The seventh chapter shows that the hierarchy is only partially able to identify the best teachers, and that in the long run, these teachers are not better paid compared to the other. This chapter also demonstrates a fairly large variation across teachers as far as efficiency is concerned, similar to the one found in other international studies, and confirms the fact that these differences in term of efficiency are only weakly related to observable variables such as gender or experience.

The last part discusses the pro-competitive effect of school liberalization policies, and more specifically, the one trying to make catchment area systems evolve toward free choice systems. The chapter 9 thus presents a model of competition among schools,

which highlights the importance of the role of the peers in the assessment of school quality on the competition between schools. Indeed, when families give much weight to peers, and that the school system is stratified between more or less prestigious schools, then the schools which are gathering the best students, at the top of the hierarchy, gain a strong market power which reduce their incentive to exert efforts. By increasing vertical differentiation across schools, a policy aiming at increasing school choice can therefore lead to a decrease in the competition across schools, and reduce the level of effort.

The original works presented in this dissertation, as well as the discussion of the existing literature on the topic that it does enable to conclude that market mechanisms are not an effective solution to the problems of performance and student under-achievement that education systems in developed countries have to face. Indeed, the peculiarities of the market for education, where peer effects, imperfect information, and bounded rationality are present, make public education systems as efficient as, and maybe more efficient than private systems.

However, as it is shown in the third part, public education systems have to face agency problems, and have difficulties to manage and reward their staff in a satisfactory way. As a result, it is needed to find alternative mechanism enabling to improve the efficiency of the supply of education.





# Table des matières

<b>Remerciements</b>	<b>5</b>
<b>Résumé</b>	<b>9</b>
<b>Table des matières</b>	<b>17</b>
<b>I Introduction</b>	<b>23</b>
1 Motivation du sujet . . . . .	24
2 Les politiques de libéralisation scolaire en pratique . . . . .	26
3 Les avantages attendus de la mise en place d'un marché de l'éducation .	30
4 Une vision trop idéale ? Les obstacles à la mise en œuvre d'un marché de l'éducation . . . . .	32
5 Plan de thèse . . . . .	34
<b>II Demande d'éducation et choix d'orientation</b>	<b>39</b>
<b>1 La vision économique du choix scolaire et de l'orientation</b>	<b>41</b>
1.1 Introduction . . . . .	42
1.2 Définir le choix scolaire . . . . .	43
1.3 Préférences et rationalité des parents d'élèves . . . . .	44
1.4 Les établissements peuvent-ils sélectionner correctement leurs élèves ? .	45
1.5 L'allocation des élèves aux écoles dans un contexte de rareté . . . . .	47
1.6 Conclusion . . . . .	48
<b>2 Matching with endogenously revealed preferences</b>	<b>51</b>
2.1 Introduction . . . . .	52
2.2 Some stylised facts : the distribution of the size of school markets and the existence of peer effects . . . . .	55
2.2.1 The distribution of school market sizes . . . . .	55
2.2.2 Review of the evidence on the existence of peer effects . . . . .	56
2.3 A model of formation of school preferences . . . . .	57

2.4	Numerical and theoretical results : the impact of preference for peers on welfare . . . . .	59
2.5	conclusion . . . . .	64
2.6	Appendix . . . . .	65
<b>3</b>	<b>Mismatch or convergence ? Affirmative action and students' selection in a French elite university</b>	<b>69</b>
3.1	Introduction . . . . .	70
3.2	Data and institutional environment . . . . .	73
3.2.1	The affirmative action program . . . . .	73
3.2.2	Data . . . . .	74
3.3	Do affirmative action students catch up? . . . . .	77
3.3.1	Identification strategy . . . . .	77
3.3.2	The academic achievement of affirmative action students . . . . .	79
3.4	Is the selection process rational? . . . . .	85
3.4.1	The determinants of failure and convergence . . . . .	85
3.4.2	The impact of affirmative action students on their peers . . . . .	87
3.5	Conclusion . . . . .	88
3.6	Appendix . . . . .	90
3.6.1	Appendix I : the allocation of affirmative action students across groups . . . . .	90
<b>III</b>	<b>Les effets de pairs</b>	<b>93</b>
<b>4</b>	<b>L'économétrie des réseaux et des effets de pair</b>	<b>95</b>
4.1	Introduction : l'importance des effets de pair pour un marché scolaire . . . . .	96
4.2	Effets de pairs et problèmes d'identification . . . . .	97
4.3	Les méthodes classiques : expériences naturelles, variables instrumentales et effets fixes . . . . .	99
4.4	L'approche par les effets fixes . . . . .	103
4.5	Effets de pair et réseaux . . . . .	104
4.6	Les modèles structurels de réseaux . . . . .	108
4.7	Estimer les effets de réseaux sans connaître la matrice d'interaction . . . . .	109
4.8	Conclusion . . . . .	111
<b>5</b>	<b>Measuring endogenous and exogenous peer effects using class size variation</b>	<b>113</b>
5.1	Introduction . . . . .	114
5.2	Data and institutional context . . . . .	116
5.3	Identification strategy . . . . .	118
5.3.1	The econometric model . . . . .	118

5.3.2	Missing data problem . . . . .	119
5.4	Estimation and results . . . . .	120
5.5	Conclusion . . . . .	125
5.6	Appendix 1 : Detail of the procedure to maximize the log-likelihood . .	126
<b>IV</b>	<b>Le marché du travail des professeurs</b>	<b>129</b>
<b>6</b>	<b>La gestion des ressources humaines dans le secteur éducatif</b>	<b>131</b>
6.1	introduction . . . . .	132
6.2	Les différences d'efficacité entre enseignants . . . . .	134
6.3	Le recrutement et la formation des enseignants . . . . .	137
6.4	Les dispositifs d'incitation des professeurs . . . . .	143
6.5	L'appariement entre élèves et enseignants . . . . .	147
6.6	conclusion . . . . .	149
<b>7</b>	<b>Does teachers' compensation reward effectiveness</b>	<b>153</b>
7.1	Introduction . . . . .	154
7.2	Data and descriptive statistics . . . . .	158
7.2.1	Students . . . . .	158
7.2.2	Teachers . . . . .	160
7.3	Identification strategy and results . . . . .	163
7.3.1	Identification strategy . . . . .	163
7.3.2	Results on the link between pay and efficiency . . . . .	163
7.4	Measuring teachers' effects dispersion . . . . .	168
7.4.1	The first stage equation . . . . .	169
7.4.2	Estimating the variance . . . . .	170
7.4.3	Computing the mean teacher effect . . . . .	171
7.4.4	The variance of teacher effects . . . . .	172
7.5	Conclusion . . . . .	174
7.6	Appendix . . . . .	175
7.6.1	Table 4 : specification with no fixed effects . . . . .	175
7.6.2	Table 4 : restricted samples with only teachers from 6 to 20 years of experience . . . . .	176
<b>V</b>	<b>La concurrence entre offreurs d'éducation</b>	<b>181</b>
<b>8</b>	<b>Le choix scolaire : un moyen d'améliorer l'efficacité du système éducatif?</b>	<b>183</b>
8.1	Introduction . . . . .	184
8.2	Les arguments en faveur d'un effet pro-compétitif du choix scolaire . . .	185
8.3	Des résultats empiriques mitigés . . . . .	187

8.4	Les nouvelles théorisations du marché éducatif : stratification et information imparfaite . . . . .	189
8.5	Conclusion . . . . .	191
<b>9</b>	<b>School competition and endogenous vertical differentiation</b>	<b>195</b>
9.1	Introduction . . . . .	196
9.2	Previous models of school competition . . . . .	200
9.3	A model of imperfect competition between non-profit schools . . . . .	202
9.3.1	General model . . . . .	202
9.3.2	Resolution with two schools, a continuum of ability type, and explicit functional forms . . . . .	205
9.4	Numerical results and comparative statics . . . . .	207
9.4.1	Existence of an equilibrium . . . . .	207
9.4.2	The impact of peer effects on optimal effort : numerical results . . . . .	208
9.5	Discussion and empirical prediction . . . . .	211
9.6	Conclusion . . . . .	212
9.7	Appendix . . . . .	212
9.7.1	Values used in the numerical simulation . . . . .	212
9.7.2	Method of resolution of the model . . . . .	213
<b>VI</b>	<b>Conclusion</b>	<b>217</b>
1	Une tentative de bilan . . . . .	218
2	Implication pour les politiques publiques et pistes de recherches futures . . . . .	219
	<b>Bibliographie</b>	<b>223</b>





# Première partie

## Introduction

# 1 Motivation du sujet

Bien qu'elle soit d'une importance fondamentale pour toute société, l'éducation n'a attiré l'attention des économistes que récemment, notamment avec les travaux de Becker (2009) [22], Mincer (1958) [130] et Schultz (1961) [154] sur le capital humain. En effet, dans la plupart des pays développés, l'éducation est le lieu d'une intervention massive de l'État, qui a rendu l'instruction primaire et secondaire obligatoire depuis la fin du XIX<sup>ème</sup> siècle. Quelque soit le degré de décentralisation du système éducatif, il s'agit la plupart du temps des pouvoirs publics qui font respecter l'obligation de scolarisation, financent les dépenses éducatives, et déterminent les programmes. Même lorsqu'un secteur privé existe, il est en général minoritaire et fortement régulé.

La France est un bon exemple de cette intervention de l'État à tous les niveaux de la sphère éducative. C'est ainsi le ministère de l'éducation nationale qui organise les concours de recrutement des enseignants, qu'ils aient à vocation à enseigner dans le secteur public ou privé, qui fixe au niveau national les programmes obligatoires pour l'enseignement public comme privé, et qui assure l'essentiel du financement de l'éducation, au point que 80% des dépenses du secteur privé sont financées en réalité par l'État. Bien qu'extrême, cette situation n'est pas exceptionnelle. Même des pays réputés libéraux comme les États-Unis se caractérisent par une très forte intervention publique.

Située en dehors du marché, la sphère éducative semblait donc d'un intérêt limité pour les économistes. Elle paraissait avant tout relever du champ d'étude des pédagogues, des sociologues de l'éducation, et des psychologues de l'apprentissage. Cependant, depuis les années 1960, un grand nombre de travaux d'économiste portant sur des questions éducatives se sont développés, faisant de l'économie de l'éducation un champ dynamique et en expansion. Deux grands axes ont été poursuivis.

Tout d'abord, la science économique s'est intéressée à l'éducation du point de vue de la demande. L'éducation peut en effet être considérée comme un bien d'investissement, susceptible d'apporter un rendement en augmentant les salaires futurs. Cette observation a motivé une littérature considérable sur les rendements de l'éducation, cherchant à déterminer la rentabilité de l'investissement éducatif, aussi bien pour différents niveaux d'étude que pour différents types d'individu. Les implications de ces travaux étaient claires : il s'agissait de déterminer dans quelle mesure il était intéressant d'investir dans l'éducation, et jusqu'à quel point.

Un second axe s'est, logiquement, intéressé à l'offre d'éducation, en cherchant à mesurer l'influence des différents facteurs de production éducatif dans l'apprentissage, ainsi que les facteurs influençant la qualité des écoles. Cet axe de recherche a notamment donné lieu à des travaux mesurant l'effet de la taille des classes sur l'apprentissage, les effets de pairs, ou encore les différences de capacité entre enseignants. De manière plus normative, certains de ces travaux ont également essayé d'évaluer la possibilité de mise en œuvre de dispositifs d'incitation inspirés de la théorie des contrats et de l'agence.

Bien qu'ayant généré de nombreux travaux très intéressants, ces deux axes relèvent cependant plus des méthodes quantitatives que de l'économie proprement dite. En effet, il s'agit avant tout de problèmes de mesure, qui sont traités majoritairement par les économistes en raison de l'avantage comparatif qu'ont ceux-ci par rapport aux scientifiques issus d'autres sciences sociales, peu formés aux méthodes quantitatives.

Une autre angle d'approche, proprement économique, des questions éducatives consistent à se poser la question de la place respective de l'État et du marché dans l'éducation. La répartition des rôles entre intervention publique et mécanismes de marché est une question classique que se posent les économistes, et la réponse la plus fréquemment apportée est que l'État doit avant tout se charger des biens publics, c'est à dire des biens qui ne sont ni rivaux (leur consommation ne se fait pas au détriment de la consommation d'autrui), ni exclusif (il est impossible d'empêcher quelqu'un de consommer ces biens). Ce type de bien est typiquement produit en quantité insuffisante par le marché, ce qui justifie une production par l'État (ou du moins, une production financée publiquement).

Cependant, l'éducation ne semble pas faire partie de la catégorie des biens publics. Les ressources éducatives sont rares et ne peuvent être aisément partagées (un enseignant ne peut enseigner qu'à un nombre limité d'élèves, par exemple), et il est facile de réserver l'usage des ressources éducatives à des individus précis (comme le montre l'existence d'un enseignement privé). Le poids important du secteur public dans le secteur éducatif n'a donc rien d'une évidence d'un point de vue économique.

Cette observation a motivé à partir des années 1970 un important débat dans les pays anglo-saxons sur la nécessité d'introduire des mécanismes de marché dans le secteur éducatif. Le partisan le plus célèbre d'une telle introduction est le prix Nobel d'économie Milton Friedman (1955) [69]. Friedman fait remarquer que l'organisation la plus courante des systèmes éducatifs publics consistent à définir des secteurs géographiques correspondants à des écoles, et à affecter tous les élèves d'un secteur à l'école correspondante. Une telle organisation revient à donner un monopole de fait aux écoles sur tous les élèves habitant dans leur secteur administratif. Or, les monopoles sont généralement caractérisés par leur inefficacité. Ne faisant face à aucune incitation de la part de concurrents, ils innovent peu, et extraient une rente au détriment des consommateurs (qui sont ici les élèves et leurs parents).

Partant de ce constat, Friedman propose donc une solution simple susceptible d'augmenter sans coût l'efficacité du système éducatif. Plutôt que de financer directement les écoles, il propose de distribuer à chaque famille un "chèque-éducation" (*voucher*) pouvant être dépensé dans n'importe quelle institution éducative, publique ou privée, non-lucrative ou marchande. Cela revient donc à subventionner la demande d'éducation, et à déréguler complètement l'offre. Le système éducatif devient alors organisé comme n'importe quel marché.

Au delà des chèques-éducation, d'autres propositions ont été faites afin de rendre le

marché éducatif plus compétitif. Ces propositions se concentrent sur l'élargissement du choix laissé à la famille dans la détermination de l'école de leur enfant. Beaucoup de systèmes publics ont ainsi mis fin à la sectorisation, en donnant la possibilité à des familles résidant en dehors du district de recrutement d'une école de demander la scolarisation de leur enfant dans l'établissement. Certains pays ont également cherché à stimuler l'entrée de nouveaux acteurs sur le marché, en favorisant les groupes de parent désireux de fonder des écoles alternatives pour éduquer leurs enfants (c'est l'exemple des *charter schools* aux États-Unis).

Ainsi, en quelques décennies, les mécanismes de marché se sont répandus dans de nombreux systèmes éducatifs, même si la place de l'État reste prééminente. Cette expansion du marché renouvelle donc l'intérêt des économistes pour l'éducation. Il ne s'agit plus seulement de mesurer les facteurs déterminant l'offre et la demande d'éducation, mais aussi d'analyser comment les marchés éducatifs fonctionnent, et s'ils apportent un avantage réel par rapport à l'intervention publique. Cette dernière question sera donc le fil directeur de la thèse :

**Dans quelle mesure les mécanismes de marché permettent-ils d'améliorer l'efficacité du système éducatif par rapport à un système public centralisé ?**

Bien évidemment, il n'est pas possible de répondre entièrement à cette question dans le cadre d'une thèse de doctorat. La thèse s'attachera donc à présenter une série d'essais apportant des éléments de réponses à cette question générale, aussi bien à un niveau théorique qu'empirique. Par ailleurs, autant que possible, la thèse essaiera de mettre en perspective ces essais avec la littérature scientifique existante, afin de donner une perspective plus large au lecteur sur la question. L'ambition de la thèse est donc de se positionner entre le modèle de la thèse par article, et celui de la thèse conçue comme un ouvrage unique et cohérent.

## 2 Les politiques de libéralisation scolaire en pratique

Si l'introduction de mécanismes de marché<sup>1</sup> et le rapprochement avec le fonctionnement du secteur privé ont été des idées populaires ces trente dernières années, leur mise en œuvre a pris des formes variées. En effet, pour des raisons politiques, la plupart des systèmes scolaires ont été réticents à mettre complètement fin à toute intervention de l'État.

Ces réticences ont été particulièrement importante dans le domaine du financement.

---

1. On utilisera dans la suite du texte de manière interchangeable les expressions "introduction de mécanismes de marché", "dérégulation" et "libéralisation"

En pratique, aucun système scolaire n'a renoncé à un financement public de l'éducation pour d'évidente raison d'égalité entre enfants et de protection de leurs droits fondamentaux. Dans sa version la plus libérale, le financement se fait à travers l'attribution de subventions aux familles, les écoles restant libres de demander des frais supplémentaires. Un tel système peut donc générer des inégalités non négligeables, dans la mesure où certaines écoles peuvent se spécialiser en ciblant les familles les plus riches prêtes à payer beaucoup pour obtenir une éducation de qualité supérieure. Par conséquent, au sein du groupe des pays à revenus élevés ou intermédiaires, il n'a été mis en œuvre qu'au Chili, où des écoles à prix variés peuvent coexister, desservant différents groupes de la population.

La plupart des systèmes scolaires mettant en œuvre des politiques de libéralisation ont préféré interdire aux écoles de demander des frais additionnels aux parents, ou bien ont strictement encadré le montant de ces frais. Cela revient donc à une forme de contrôle des prix. Les écoles entrent alors en concurrence uniquement afin d'attirer un plus grand nombre d'élève, et donc attirer des subventions supplémentaires. Ce système a parfois été ajusté pour tenir compte du fait que tous les étudiants n'ont pas le même coût. L'État accordent alors des subsides différenciés aux écoles en fonction du type d'élève accueilli, avec des bonus pour l'accueil, par exemple, d'élèves handicapés ou souffrant de difficultés spécifiques d'apprentissage.

Un deuxième aspect provoquant des réticences a été l'introduction d'écoles à but lucratif. Les théories de la firme insistent en effet sur l'importance du caractère lucratif des entreprises pour expliquer leur succès relatif par rapport à d'autres formes d'organisation de la production (voir par exemple Alchian et Demsetz, 1972 [6], et Jensen et Meckling, 1976 [104]). En attribuant à un type d'agent, les actionnaires, tous les bénéfices restant après que les coûts de production ont été couverts, ainsi que le pouvoir de décision, les entreprises mettent en place un système d'incitation efficace. Ce sont ceux qui ont le plus à perdre d'une mauvaise gestion qui ont le pouvoir, et qui sont les "*residual claimants*". Par conséquent, les bénéfices à attendre d'un marché de l'éducation ne sont pas uniquement dus à l'introduction de concurrents, mais aussi à une volonté des écoles de gagner des parts de marché en raison de leur volonté de faire du profit.

Cependant, les caractéristiques spécifiques de l'éducation font qu'il est difficile de l'organiser uniquement à travers des organisations à but lucratif. En effet, les services éducatifs sont difficilement contractualisables : une école ne peut s'engager sur la qualité de l'éducation qu'elle fournit. De plus dès lors qu'un élève est scolarisé dans une école, il est difficile d'en changer avant plusieurs années, ce qui réduit la capacité des consommateurs d'exprimer leur insatisfaction. Enfin, il existe une asymétrie en terme de rapport de force entre l'école et les familles, dans la mesure où l'école reste une organisation hiérarchique où les élèves sont des individus soumis à une autorité avant d'être des consommateurs. Pour toutes ces raisons, la plupart des familles sont réticentes à faire appel à des fournisseurs d'éducation à but lucratif, dans la mesure où les

problèmes d'aléa moral sont larges, et qu'il semble y avoir une contradiction entre l'accomplissement d'une mission éducative nécessitant désintéressement, voire altruisme, et volonté de faire du profit. Ces raisons expliquent également pourquoi certains systèmes scolaires interdisent explicitement l'entrée d'acteur à but lucratif sur le marché de l'éducation.

La grande majorité des institutions éducatives sont donc des organisations non-marchande, et sont souvent organisées de manière coopérative ou associative. Dans ce cas, le pouvoir au sein de l'organisation est exercé par un conseil de parents, dont les intérêts coïncident parfaitement avec la recherche de l'excellence éducative à moindre coût. Cependant, des problèmes d'agence peuvent subsister, dans la mesure où les parents (de même que les actionnaires), font face à un problème d'action collective. Si le contrôle du personnel enseignant apporte des bénéfices à tous les élèves, il a un coût individuel. Par conséquent, chaque famille a intérêt à laisser aux autres le soin d'exercer ce contrôle. Dans le cas des firmes, d'autres mécanismes, tels que les LBO et les OPA/OPE existent, permettant de donner de bonnes incitations au principal (1986) [103]. De tels mécanismes sont inexistants dans le cas des écoles. En pratique, le directeur d'une école cherche sans doute plus à trouver un équilibre entre les exigences des enseignants et celles des parents d'élèves que de maximiser la qualité éducative de son établissement.

On le voit l'analogie entre l'école et l'entreprise d'une part, et un marché de l'éducation et un marché ordinaire d'un bien ou d'un service d'autre part est très loin d'être parfaite. Le marché de l'éducation est en pratique un marché où les prix sont contrôlés et où les acteurs sont des institutions à but non lucratifs, non cotées, et vulnérables aux problèmes d'agence. Les expériences poussant jusqu'au bout l'analogie et encourageant l'entrée d'acteurs lucratifs sont très rares. En l'espèce, seul le Chili a pleinement essayé l'usage de mécanismes de marché dans le système éducatif. La plupart des autres pays désireux d'encourager la concurrence ont en réalité promu avant tout le libre choix scolaire, sans remettre en question la dimension non-lucrative de la sphère éducative.

Un autre aspect faisant débat, et souvent l'objet d'une régulation, concerne les politiques d'admission et d'allocation des élèves aux écoles. En effet, le choix scolaire ne s'exerce pas sans contrainte, dans la mesure où la plupart des écoles font face à des contraintes de taille, et essaient fréquemment d'attirer certains élèves plutôt que d'autres. En particulier, dans la mesure où l'attractivité d'une école dépend de manière importante de la qualité des élèves qui la fréquente, beaucoup d'établissements ne souhaitent pas admettre les élèves les plus faibles, qui risquent de dégrader leur réputation. Laisser aux établissements la liberté de sélectionner leurs élèves aboutit donc à une stratification scolaire, où il existe une hiérarchie d'établissement se distinguant par la qualité de leurs élèves, plus que par leur efficacité intrinsèque.

Désireux d'éviter un tel résultat, certains systèmes scolaires ont donc fait le choix de restreindre la capacité des écoles à sélectionner leur élèves. En cas de surcapacité, les élèves sont alors choisis en suivant un algorithme qui recourt largement au hasard pour

attribuer des places à des élèves ayant le même niveau de priorité dans une école. On s'éloigne donc encore plus du modèle du marché des biens non régulé : les écoles n'ont même plus la capacité de choisir leurs consommateurs.

Enfin, une dernière dimension éloignant le marché de l'éducation d'un marché traditionnel a trait à l'autonomie laissée aux établissements. En effet, tous les systèmes éducatifs imposent des obligations aux établissements scolaires, que cela soit en terme de contenu des cours, de recrutement des enseignants, voire d'organisation interne. Les politiques de libéralisations peuvent donc différer dans leur volonté de contrôler plus ou moins étroitement les établissements scolaires. Certains systèmes choisissent en effet de combiner une extension du choix des familles avec un renforcement du contrôle de l'État sur le fonctionnement des écoles. On peut donc choisir d'introduire simultanément des mécanismes de marché et des mécanismes de contrôle étatiques traditionnels (c'est par exemple le cas des Etats-Unis qui ont simultanément encouragé la création de *charter schools* et renforcé le contrôle sur les écoles subventionnées par l'Etat à travers le programme "*No child left behind*").

On le voit, plutôt que de deux modèles opposés, l'un faisant appel au marché, l'autre à l'État, il existe plutôt une continuité de modèles nationaux spécifiques allant de modèles très centralisés et étatisés, comme le modèle français, à des modèles essayant de répliquer les marchés traditionnels des biens et services, comme le Chili.

Passer en revue l'ensemble des modèles d'éducation nationaux est très au delà de la portée de cette introduction. On peut cependant citer quelques exemples de systèmes éducatifs ayant tentés de mettre en œuvre des politiques de libéralisation.

La réforme la plus achevée de libéralisation est sans aucun doute celle mise en œuvre au Chili, pendant la dictature d'Augusto Pinochet. A partir de 1981, le secteur public a été très largement décentralisé, et le budget des écoles publiques strictement lié à leur fréquentation. Un mécanisme de "*vouchers*" a été mis en place, permettant une large expansion du secteur privé, principalement à but lucratif. Les écoles privées se sont vu accorder une liberté complète de sélectionner leurs élèves, ainsi que, de manière progressive, la liberté de demander des droits d'inscription additionnels (Hsieh et Urquiola, 2006 [94], Neilson, 2014 [133]).

Réputés libéraux, les États-Unis ont conservé un système d'éducation principalement public, bien que décentralisé en districts scolaires. Plusieurs tentatives ont été mené afin d'étendre le choix des élèves, et de faire émerger une offre privé. En particulier, les "*charter schools*", des écoles privés sans but lucratif, gérées par des conseils de parents, ont été introduites et ont bénéficié de subvention spéciale. Le but de ce programme est de susciter l'entrée de nouveaux acteurs sur le marché éducatif, potentiellement plus efficace, et capable d'offrir une alternative au secteur public. Il reste cependant d'ampleur limité. Des programmes de libre choix scolaire ont également été mené dans certains districts au niveau du secondaire. Cependant, ils reposent sur un système

centralisé d'allocation des étudiants utilisant des mécanismes d'appariement, ce qui les éloignent d'un mécanisme de marché.

Les pays européens ont conservés des systèmes principalement publics, et fondé sur le principe de la sectorisation. Cependant, certains d'entre eux ont essayé d'introduire plus de liberté de choix et d'autonomie. C'est notamment le cas de la Grande-Bretagne, qui a développée la liberté de choix entre les écoles d'un même district scolaire, et encouragée l'autonomie (Gibbons et al. 2008 [76], Clark, 2009 [49]). De même, la Suède a mis en œuvre des réformes donnant une grande liberté de choix aux familles (Bjorklund, [25]). Par ailleurs, de nombreux pays conservent un système privé confessionnel subventionné opérant en même temps que le système public (la France et les Pays-Bas sont de bons exemples d'un tel système), ce qui est une forme de dérégulation encadrée.

Il est donc possible de trouver une large variété de degré de libéralisation des marchés éducatifs. Cette libéralisation répond fréquemment à des impératifs d'ordre idéologique, dans la mesure où les sociétés varient dans leur volonté de contrôler collectivement l'éducation. Cependant, on peut se demander dans quelle mesure ces politiques de libéralisations sont justifiées par des arguments liés à l'efficacité.

### **3 Les avantages attendus de la mise en place d'un marché de l'éducation**

Des arguments variés ont été avancés en faveur d'une politique de libéralisation, s'appuyant sur les effets supposés des différents aspects d'une politique de libéralisation.

L'avantage attendu le plus souvent mis en avant est l'impact pro-concurrentiel généré par l'extension de la liberté de choix des familles. On retrouve ici l'argument original de Milton Friedman, critiquant la forme monopolistique que prend la production d'éducation. En permettant aux familles de choisir librement une école, et en liant ressources et inscription, une politique de choix scolaire contraint les écoles à entrer en concurrence afin d'attirer suffisamment d'élèves pour pouvoir continuer à opérer. Une école dont la performance reste en deçà de ses concurrents sur une longue période est ainsi vouée à disparaître.

Dans cette première version de l'argument pro-compétitif du choix scolaire, celui-ci n'a donc un impact que dans la mesure où il fait peser une menace existentielle sur les écoles. Une version raffinée de l'argument insiste sur le fait que gagner des étudiants supplémentaires permet d'attirer des ressources additionnelles, susceptibles d'être redistribuées vers le personnel de l'école, où vers le propriétaire de l'école si celle-ci est à but lucratif. On voit donc ici à quel point la forme organisationnelle prise par l'école est importante, dans la mesure où une école à but non-lucratif, et gérée dans l'intérêt

exclusif des parents est moins susceptible de répondre à une pression compétitive.

Une politique de dérégulation comprend également la libéralisation de l'entrée sur le marché éducatif, et le recours au secteur privé. Cette suppression des barrières à l'entrée peut avoir des effets positifs dans la mesure où elle incite de nouveaux acteurs plus efficaces à opérer sur le marché éducatif. Elle nécessite donc une plus grande efficacité du secteur privé par rapport au secteur public existant.

Un autre canal par laquelle une telle politique est susceptible d'améliorer l'efficacité du système éducatif réside dans les possibilités de ré-allocations qu'elle crée entre acteurs. En effet, elle donne la possibilité aux acteurs les plus efficaces d'ouvrir de nouveaux établissements dans d'autres zones géographiques, et donc de remplacer progressivement les acteurs les moins performants. De tels bénéfices supposent donc une hétérogénéité importante en terme d'efficacité entre acteurs éducatifs, et surtout la capacité de dupliquer les meilleures formes organisationnelles.

Si la littérature sur les différences de productivité entre firmes montre que ces conditions sont bien remplies dans le cas des marchés traditionnels (Bloom et Van Reenen, 2007 [27], Hsieh et Klenow, 2009 [93] ), cela reste cependant à vérifier pour le marché éducatif.

Un troisième grand type d'avantage souligné par les partisans d'une politique de dérégulation scolaire est l'impact positif d'une plus grande autonomie des écoles sur leur efficacité. En effet, les systèmes publics sont souvent centralisés, soit directement en raison d'un contrôle étatique étroit sur l'ensemble du système, soit indirectement en raison de la présence de syndicats d'enseignants puissants capables de négocier des conventions collectives couvrant l'ensemble des écoles et des professeurs. Une telle centralisation rend plus difficile la résolution des problèmes d'agence concernant les enseignants, ainsi que l'adaptation de la gestion des écoles aux impératifs locaux.

Or, la plupart des politiques de libéralisation s'accompagne d'un transfert de pouvoir vers l'échelon local. Ainsi, la réforme chilienne a impliqué l'abrogation des conventions syndicales régissant la gestion du corps enseignant. Une plus grande décentralisation du pouvoir doit donc permettre aux directeurs d'écoles de gérer plus finement leurs ressources humaines, et surtout de licencier les enseignants les plus incompetents et donner de meilleures incitations aux autres. L'autonomie est donc supposé promouvoir une gestion plus efficace des ressources au niveau local.

L'autonomie est également complémentaire avec les incitations données par le choix scolaire et la libre entrée. En effet, l'autonomie sans incitations à s'améliorer aggrave plutôt les problèmes d'inefficacité des écoles, en augmentant les possibilités de mauvaises gestions, tandis que de meilleures incitations auxquelles les agents ne peuvent réagir sont vouées à rester sans effet.

Enfin, un dernier argument, plus marginal, avancé en faveur de la libéralisation scolaire est sa capacité à élargir la variété de l'offre éducative proposée aux familles, et à mettre en œuvre un meilleur appariement entre écoles et élèves. En effet, les familles

peuvent être à la recherche d’une expérience éducative différente pour leurs enfants, ou bien certains enfants peuvent avoir des besoins spéciaux mal couverts par les écoles traditionnelles. Si la dérégulation porte aussi sur le contenu des programmes et l’organisation du cursus, elle est susceptible d’apporter un gain important en permettant aux écoles de se différencier horizontalement et de se spécialiser dans l’accueil d’un certain type d’élève. Ces gains se rapprochent de ceux mis en avant dans la nouvelle littérature sur le commerce international, qui montre comment l’ouverture des marchés met à disposition de nouveaux biens et services auprès des consommateurs.

Cependant, cet argument a au moins autant une dimension philosophique que économique. En effet, pour beaucoup, le fait de délivrer à tous une éducation homogène est une fin en elle-même, et vise à promouvoir un sentiment d’appartenance à une société commune. L’existence d’un gain lié à la variété dépend donc d’une opinion philosophique sous-jacente sur ce que doit être la finalité de l’éducation. Il a cependant un contenu économique susceptible d’être évalué : en effet, ce gain ne peut exister que si un marché dérégulé permet un meilleur appariement entre élèves et écoles. Or, en présence de frictions importantes, la capacité d’un marché libre à atteindre un tel objectif peut être mise en question.

Les partisans d’un marché scolaire avancent donc des arguments nombreux et cohérents en faveur d’une libéralisation. Cependant, ces arguments reposent sur une analogie entre le marché scolaire et les marchés traditionnels des biens et services qui est discutable. Comme va essayer de le montrer la section suivante, le marché de l’éducation comprend des caractéristiques le rendant très singulier, et nécessitant de développer une analyse économique spécifique.

## 4 Une vision trop idéale ? Les obstacles à la mise en œuvre d’un marché de l’éducation

Le premier point majeur sur lequel le marché éducatif diffère des marchés traditionnels est la contribution des consommateurs eux-mêmes à la qualité du produit qu’il consomme. En effet, l’école n’est pas seulement un lieu d’éducation, elle est aussi un lieu de socialisation, ayant une influence sur de nombreuses capacités non cognitives. Elle est aussi le lieu où se forme la plupart des relations amicales des enfants.

Par conséquent, la qualité d’une école dépend de manière cruciale des élèves qui la fréquentent, et qui sont susceptibles d’exercer une influence sur les autres. Cette influence est appelé dans la littérature économique et sociologique “effets de pairs”. Empiriquement, si les effets de pairs cognitifs sont faibles, les effets de pairs portant sur des comportement non-cognitifs (tels que la délinquance, les comportements à risque, l’usage de drogue) sont importants (Epple et Romano, 2011 [58]). Par ailleurs, quelque soit la réalité des effets de pairs, il existe de nombreuses preuves empiriques montrant

que les familles accordent une grande importance à la composition sociale et ethniques des écoles où sont scolarisés leurs enfants.

Cette caractéristique du marché éducatif donne lieu à une différenciation verticale entre école nommée stratification (Epple et Romano, 1998 [60]). Les meilleurs élèves tendent à se regrouper ensemble, et il existe une hiérarchie des écoles au sein du système éducatif, aussi bien au point de vue des résultats scolaires que de la composition sociale (ces deux dimensions étant fortement corrélées). Cette différenciation peut potentiellement avoir un impact négatif sur la concurrence, les écoles les plus prestigieuses possédant un pouvoir de marché, car étant très demandée.

Une autre conséquence de la présence d'effets de pair est qu'il devient optimal pour les écoles d'adopter une politique de tarifs et d'admission différenciés en fonction des élèves. Autant que possible, les écoles cherchent soit à exclure, soit à faire payer plus les élèves les plus faibles, pour prendre en compte l'externalité négative qu'ils exercent sur les autres.

Enfin, une dernière conséquence de la présence d'effets de pairs est que l'allocation des élèves au sein du système scolaire a un impact important en terme d'équité et d'inégalité. En effet, un système très stratifié est aussi un système très inégalitaire, les élèves les plus faibles et les plus défavorisés socialement recevant une éducation de moindre qualité, car se retrouvant dans les écoles avec les pairs les plus médiocres.

La présence d'effets de pairs a donc un impact important sur la structure du marché de l'éducation, et sur la stratégie optimale des écoles. En créant un pouvoir de marché, elle peut mener à des inefficacités, et à des résultats très loin de ceux du modèle idéal de concurrence pure et parfaite. Ces inefficacités sont susceptibles d'être exacerbées par le fait que la liberté des écoles de fixer leurs prix et d'admettre leurs élèves est souvent contrainte, ainsi que par le fait que les écoles ont souvent un objectif non lucratif.

Le second point sur lequel les marchés éducatifs se distinguent des marchés traditionnels est liés à la présence d'imperfection dans l'information disponible, et de rationalité limitée. La nécessité de faire un choix impose en effet de bien connaître les alternatives disponibles. Or, il n'est pas évident d'évaluer une école. En effet, les résultats d'une école dépendent aussi bien de ses pairs que de son efficacité intrinsèque, ces deux facteurs étant difficiles à distinguer. Par ailleurs, l'efficacité d'une école peut varier en fonction du type de l'élève. Enfin, le signal le plus facilement disponible, à savoir les scores à des test standardisés, ne dépend que partiellement de l'école et est avant tout influencé par la qualité des élèves.

Évaluer la qualité des écoles est donc difficile. Or, le cœur de l'argument des partisan des politiques de libéralisation est que le libre choix permet aux meilleures écoles d'attirer plus d'élève. Ce mécanisme ne peut être que considérablement atténué si les familles font leur choix de manière imparfaite.

Les problèmes d'information et de rationalité limitée ont également un impact sur la qualité de l'appariement entre école et élève, supposé être meilleur dans un système de

libre choix. Outre le fait que les familles peuvent préférer des écoles qui ne sont pas les meilleures en terme de qualité, elles peuvent également faire des erreurs en terme d'orientation dans une filière académique. Il est possible, par exemple, que des familles orientent leurs enfants vers des parcours trop difficiles, les exposant à l'échec scolaire, ou bien qu'à l'inverse, elles manquent d'ambition et ne choisissent pas les filières les plus rentables sur le long terme pour leurs enfants. Ces erreurs peuvent également être commises par les établissements éducatifs eux-même, qui peuvent se tromper sur le potentiel de réussite d'un élève lors de la décision d'admission.

Enfin, une autre difficulté est lié à l'existence de frictions de recherche sur le marché éducatif. En effet, il peut être couteux de faire une demande d'admission dans un grand nombre d'établissement. Le fait que certains étudiants soient admis dans plusieurs établissement simultanément, et que ces derniers doivent gérer des files d'attente peut également créer un problème dans l'allocation des étudiants aux écoles.

Il n'est pas évident qu'un système décentralisé soit plus efficace qu'un système public centralisé face à ce type de problèmes. Par ailleurs, un système de marché peut limiter les incitations à la révélation d'information, en créant des problèmes de sélections adverses. Ainsi, une école en quête d'élèves n'a guère intérêt à signaler à une famille que son programme est peu adapté aux capacités intellectuelles de leur enfant, par exemple. La présence de problèmes d'information sévères éloignent donc encore plus les marchés éducatifs du modèle de la concurrence pure et parfaite, ou même du modèle du marché traditionnel des biens, raisonnablement compétitif, qui est la référence implicite des partisans d'une libéralisation du marché.

En conclusion, les marchés éducatifs sont particulièrement complexes, et marqués par de potentielles inefficacités. Il est possible que ces inefficacités soient suffisamment larges pour que un système public centralisé soit supérieur à un système de marché. Par conséquent, les bénéfices attendues d'une libéralisation du marché n'ont rien d'une évidence, et doivent être évalués rigoureusement empiriquement et théoriquement.

## 5 Plan de thèse

La thèse essaie d'évaluer les effets d'une politique de libéralisation du marché scolaire en quatre grandes parties.

La seconde partie discute l'impact d'une libéralisation scolaire sur l'orientation.

Le chapitre 1 discute de la manière dont les économistes envisagent le choix scolaire. Il montre que plutôt qu'une opposition entre la contrainte étatique et le choix offert par le marché, il existe plutôt une variété de modèle d'allocation des élèves aux écoles, qui peuvent notamment se fonder sur le marché résidentiel, un marché scolaire décentralisé, ou encore des algorithmes centralisés d'appariement. Une question cruciale est

alors la rationalité des préférences exprimées par les familles et les écoles. En pratique, des études empiriques donnent de bonnes raisons de douter que ces préférences soient rationnelles. Une certaine forme de paternalisme, et donc d'intervention étatique, pourrait alors être justifiée.

Le chapitre 2 apporte une contribution théorique à ce débat en développant un modèle d'appariement entre élèves et écoles en présence de préférences pour les pairs de la part des ménages. Cette modélisation est originale dans la mesure où la littérature sur les mécanismes d'appariement entre élèves et écoles a presque complètement ignoré cet aspect spécifique des préférences des familles. Le chapitre montre par des simulations numériques que la présence de telles préférences pour les pairs mène à une perte d'efficacité en terme de qualité des appariement, l'efficacité étant ici comprise comme la maximisation de la somme des utilités individuelles. En effet, la présence de préférences pour les pairs, sous certaines conditions, crée une corrélation accrue entre les choix soumis par les familles aux écoles. Une telle corrélation dégrade alors la qualité de l'appariement, dans la mesure où elle rend ces choix moins informatifs sur les préférences sous-jacente pour le planificateur social. Le chapitre propose une simple modification des mécanismes actuels qui peut rendre certains d'entre eux nettement plus efficaces. Il montre aussi qu'il existe un bénéfice social à une ségrégation scolaire plus faible, lié à une meilleure possibilité de satisfaire les préférences idiosyncrasiques des familles pour les écoles.

Le chapitre 3 s'intéresse à un programme de discrimination positive mis en place dans une grande école française. Il exploite l'allocation aléatoire des étudiants entre groupes en premier cycle, ainsi que l'existence d'un large tronc commun pour comparer la progression des élèves au fil du temps, et évaluer la capacité de rattrapage des élèves admis par le programme de discrimination positive. Le chapitre montre que cette convergence, bien que réelle, reste très incomplète. De plus, il existe un risque d'échec beaucoup plus élevé en premier cycle pour les élèves admis de manière préférentielle. Par contre, ces élèves ne semblent pas choisir de majeure moins lucrative ou s'orienter vers les majeures les plus aisées académiquement. Le chapitre met ensuite en évidence que des variables prédisant l'échec existent, mais sont apparemment ignorées aussi bien par l'établissement que par les élèves. L'analyse de ce programme, au delà de son intérêt intrinsèque, permet donc d'apporter un nouvel exemple de la rationalité limitée dont peuvent faire preuve élèves et établissement lors des décisions d'admission.

La troisième partie est consacrée à l'existence des effets de pairs, qui ont un impact important sur le fonctionnement d'un marché de l'éducation.

Le chapitre 4 présente une revue de la littérature économétrique (aussi bien théorique que appliquée) sur l'analyse des effets de pairs et des effets réseaux. L'analyse de la littérature montre que les effets de pairs cognitifs sont relativement faibles, mais que les effets de pairs portant sur les comportements sociaux, et notamment sur la délinquance, sont relativement élevés. Encore embryonnaire, la littérature sur les effets réseaux trouve également des effets assez faibles.

Le chapitre 5 vise à estimer les effets de pairs, aussi bien endogènes que exogènes, dans les collèges français, en utilisant une méthode récente fondée sur la variation de la taille des classes. Cette méthode est bien adaptée à ce cas dans la mesure où elle n'exige pas de variable instrumentale ou d'expérience naturelle susceptible de faire varier de manière aléatoire la composition des groupes de pairs. Or, dans les données utilisées, la ségrégation entre élèves est forte, et la corrélation entre résultats individuels et résultats du groupe est probablement due avant tout à la présence d'inobservable. La méthode utilisée consiste à comparer des élèves au sein d'une même classe, en exploitant le fait qu'au sein d'un même groupe, chaque individu fait face à des pairs légèrement différents dans la mesure où il ne s'influence pas lui-même. Le chapitre met en évidence de manière robuste l'existence d'effets négatifs associés à la présence de redoublant dans sa classe, qui sont de magnitude assez large. Il permet par contre de rejeter l'existence d'effet endogène de magnitude moyenne à large. Il est plus difficile de trancher quant à la présence d'effets de pairs concernant d'autres caractéristiques, dans la mesure où les effets sont estimés de manière très imprécises.

La quatrième partie est consacrée à la gestion des enseignants.

Le chapitre 6 présente et discute la littérature économique récente mais désormais abondante sur les enseignants. Il expose tout d'abord les travaux mesurant la variance en terme d'efficacité entre enseignants, qui montrent que la qualité des professeurs a un effet important sur le long terme, et varie très largement. Il discute ensuite la littérature sur le recrutement, la formation, et l'incitation des enseignants, et montrent qu'il est difficile d'évaluer précisément la qualité des enseignants, et que cela limite donc la possibilité de mettre en place des mécanismes d'incitation salariaux. De plus, la qualité des enseignants semble être un trait assez largement fixé, et peu modifiable par des intervention extérieures. Cependant, il semble que les directeurs d'écoles puissent dans une certaine mesure évaluer la qualité des enseignants, de manière qualitative. Les écoles semblent donc avoir en partie la capacité de contrôler le travail de leur personnel.

Le chapitre 7 analyse la politique salariale dans un système public (l'ensemble des collèges publics français), et s'intéresse à la corrélation entre salaires et efficacité des enseignants. Il contrôle pour l'allocation non aléatoire des élèves aux enseignants en utilisant des effets fixes élèves, qui permettent de comparer pour un même élève les résultats dans différentes matières enseignées par différents professeurs ayant des caractéristiques différentes. L'article montre que la hiérarchie semble disposer d'informations sur la capacité des enseignants, mais qu'elle n'utilise pas cette information pleinement pour rémunérer les enseignants au mérite. En effet, si être promu à un rythme plus rapide à un moment donné est bien associé à une plus grande efficacité, il n'y a pas de lien clair sur le long terme entre capacité à enseigner et salaire. Le chapitre mesure également les différences d'efficacité entre enseignants, et montrent qu'elles sont larges en France, de magnitude similaire à ce qui a été observé aux États-Unis. Le chapitre confirme donc l'idée que les systèmes publics centralisés d'éducation sont marqués par

une faible capacité à gérer au mieux leur ressources, et que les gains liés à une plus grande autonomie sont réels. Cependant, il ne permet pas de conclure sur une éventuelle plus grande performance du secteur privé.

La dernière partie discute de l'impact concurrentiel d'une libéralisation du marché scolaire, qui est au cœur des arguments des partisans du choix scolaire.

Le chapitre 8 discute la littérature empirique et théorique sur l'impact d'une mise en concurrence des écoles. Les partisans de la libéralisation justifient leur position en se référant la plupart du temps à des modèles d'économie industrielle très simples, dont l'application au domaine de l'éducation peut laisser sceptique. Quant aux études empiriques, elles apportent des résultats mitigés, les effets d'une compétition accrue, ou plus exactement, d'un choix plus important offert aux familles ayant plutôt des effets nuls, voire négatifs sur la réussite scolaire. Cependant, des modèles théoriques plus récents et plus sophistiqués semblent pouvoir expliquer cet apparent paradoxe, en montrant comment les écoles les plus efficaces ne sont pas forcément celles qui attirent le plus les élèves, pour peu qu'elles soient peu prestigieuses, et en faisant appel à la présence d'imperfection dans l'information.

Partant de cette petite littérature théorique, le chapitre 9 essaie de développer un modèle réaliste de concurrence entre écoles, incluant des effets de pairs, et une fonction objectif non-standard dépendant de la qualité des élèves de l'école plutôt que la taille de celle-ci, et de l'effort des professeurs. Le modèle montre que l'existence d'effets de pairs crée une différenciation verticale entre les écoles de manière endogène, et qu'une telle différenciation a pour effet de fortement réduire les incitations données par la libéralisation du marché scolaire et l'extension du choix des familles. Il permet donc de faire des prédictions quand à l'impact du passage d'un système sectorisé à un système de libre choix. Dès lors que les familles valorisent fortement la qualité des pairs, une libéralisation du marché de l'éducation peut mener à une baisse du niveau d'effort et de l'efficacité du système éducatif.

En conclusion, la thèse aboutit à des conclusions plutôt mitigées sur les conséquences à attendre de l'introduction de mécanismes de marché dans le système scolaire. En raison de l'existence d'effets de pairs et de problème de rationalité limitée, une telle réforme semble peu susceptible de générer de large gains en terme d'efficacité ou de résultats scolaires. Cependant, les systèmes publics centralisés semblent également souffrir de problème d'inefficacité. Une plus grande décentralisation et une plus grande autonomie pourrait être une manière alternative d'améliorer le fonctionnement du système scolaire.



## Deuxième partie

### Demande d'éducation et choix d'orientation



# Chapitre 1

## La vision économique du choix scolaire et de l'orientation

## 1.1 Introduction

Les avantages d'un choix scolaire accru, et plus généralement, d'une libéralisation du marché scolaire sont la plupart du temps discutés avant tout du point de vue de l'amélioration de la qualité des écoles. Cependant, une dimension importante de ces avantages est liée à l'accroissement de la variété de l'offre scolaire, et de la possibilité pour les parents de chercher une école correspondant mieux aux spécificités de leurs enfants. Un choix scolaire accru est donc supposé permettre une meilleure orientation, notamment en raison du fait que les familles sont supposées être plus susceptibles que l'État de mieux connaître ce qui est bon pour leurs enfants.

Il est notable que cette supposition n'a pas le statut d'une évidence. En effet, les politiques éducatives publiques, depuis leurs origines, ont toujours eu une dimension paternaliste importante. Ainsi, les États ont progressivement rendu la scolarité obligatoire jusqu'à la fin de l'adolescence à partir du 19<sup>ème</sup> siècle. Si il est possible dans la plupart des pays du monde de scolariser ses propres enfants à domicile, cette possibilité est la plupart du temps fortement encadrée. Par ailleurs, les programmes scolaires sont en général fixés par l'État, et les écoles privés sont tenues de s'y conformer au moins en partie.

Ce paternalisme est pour une part justifié par un objectif politique et philosophique, celui d'assurer que chaque enfant reçoit une éducation conforme aux valeurs de la société, et permettant d'acquérir ce qu'elle juge être les connaissances minimales nécessaires à chaque individu. Cependant, ce paternalisme peut également se justifier de manière pragmatique, en remettant en question le jugement des familles et des élèves. Ce jugement pourrait être affecté par des biais cognitifs, et par un manque d'information. Une intervention publique pourrait alors améliorer le bien être.

Une autre limite du choix pouvant nécessiter une intervention de l'État est que les choix sont interdépendant, dans la mesure où les familles se préoccupent de la nature des camarades de leurs enfants, et que il existe des problèmes de congestion dans le système scolaire, certaines écoles étant trop demandées par rapport à leurs capacités. Un choix scolaire décentralisé peut donc poser des problèmes de coordination entre familles, qui peuvent être résolus par l'État via la mise en œuvre de mécanismes d'appariement adéquat.

Ce chapitre essaiera donc de discuter des avantages et inconvénients d'un choix scolaire décentralisé par rapport à un système plus centralisé et régulé par l'État. Dans une première partie, il s'attachera à préciser la notion de choix scolaire. Il discutera ensuite des préférences et de la rationalité des parents d'élèves, avant de discuter du problème d'allocation des élèves aux écoles dans un contexte de rareté. Enfin, il s'intéressera à l'offre, tout d'abord en discutant de la capacité des écoles à sélectionner correctement leurs élèves, et ensuite en discutant de la structuration de l'offre en grandes filières dans le système scolaire envisagé dans son ensemble.

## 1.2 Définir le choix scolaire

Le choix scolaire est en général opposé à la sectorisation, dans la mesure où celle-ci entrave la capacité de choix des élèves. En pratique, cette distinction n'est pas vraiment pertinente dans la mesure où les parents ont la possibilité de choisir l'école de leur enfant à travers le choix résidentiel. Cependant, cette capacité est contrainte en raison du prix inégal du logement en fonction de la géographie, ainsi que des contraintes de résidences auxquelles certains ménages peuvent avoir à faire face (par exemple pour des raisons professionnelles). Les prix de l'immobilier sont ainsi affectés par la qualité des écoles, excluant de fait la possibilité de choisir pour les ménages les plus pauvres. Cependant, une libéralisation du choix des écoles n'améliore pas forcément la situation. En effet, plus que d'une liberté de choix, il s'agit avant tout d'une liberté de vœux, les écoles faisant face à des contraintes de taille qui les forcent à sélectionner les élèves. De plus, beaucoup d'écoles ont des préférences sur les élèves, dans la mesure où la qualité des élèves affecte leur qualité propre.

Par conséquent, plutôt que d'une opposition entre liberté de choix et affectation contrainte des élèves aux écoles, il s'agit plutôt d'une question d'organisation du choix, de détermination des priorités entre élèves, et de révélation des préférences.

On peut dès lors distinguer trois grands modes d'allocation des élèves aux écoles. Dans un système sectorisé, l'allocation se fait à travers le choix résidentiel. Il existe donc un marché implicite des écoles, et le choix est réservé de fait aux familles les plus riches. Dans le cas du libre choix décentralisé, les familles sont libres d'émettre des vœux directement auprès des écoles de leurs choix, qui sont libres ou non des les accepter. C'est ce système qui correspond le plus à un marché dérégulé. En pratique, dans la mesure où les écoles préfèrent souvent les meilleurs élèves, cela revient fréquemment à donner le choix aux meilleurs élèves aux dépens des autres. Enfin, dans le cas du libre choix centralisé, les élèves peuvent émettre librement leurs vœux pour n'importe quelle école, préférences des élèves et préférences des écoles sont ensuite centralisées, et un algorithme procède à l'appariement entre les deux côtés du marché.

Ces trois modèles sont en réalité entremêlés dans la mesure où il existe des dérogations à la sectorisation, et où certains élèves se voient attribuer des priorités dans certaines écoles en fonction de leur lieu de résidence dans les systèmes fondés sur le choix. Il est donc nécessaire d'analyser ces différents modes d'allocation et leur combinaison afin de déterminer dans quelle mesure une intervention publique est nécessaire.

Cependant, avant de discuter des mérites respectifs des systèmes d'allocation, il est nécessaire de discuter de la rationalité des choix faits par les parents et les écoles. En effet, chercher à satisfaire au mieux des préférences qui sont incohérentes n'a guère de sens.

### 1.3 Préférences et rationalité des parents d'élèves

Il n'est pas forcément aisé de mesurer la rationalité des parents d'élèves, dans la mesure où un choix apparemment aberrant peut aussi bien être du à une erreur réelle de calcul qu'à des préférences rares mais cohérentes. Par conséquent, cette section passera d'abord en revue les articles mesurant les préférences pour les écoles, afin de pouvoir établir ensuite si il existe des cas où les familles ne semblent pas agir en fonction de celles-ci.

L'étude pionnière dans la mesure des préférences pour les écoles est celle de Sandra Black (1999) [26], qui évalue la propension à payer des parents pour les attributs des écoles en comparant le prix des maisons de part et d'autre des limites de district scolaire (qui déterminent l'allocation à une école spécifique). Elle trouve une propension à payer non négligeable d'environ 4000 \$ pour un accroissement des scores moyens de l'école de 5 %. Ces résultats sont confirmés par Kane et al. (2006) [105] qui utilisent une stratégie similaire pour estimer la valeur de la qualité des écoles dans la ville de Charlotte en Caroline du Nord.

De manière similaire, Riegg-Cellini et al. (2010) [44] utilisent une régression sur discontinuité en exploitant le fait que les dépenses scolaires californiennes dépendent de référenda pouvant être gagnés avec une marge étroite, et trouve également un effet large sur les prix de l'immobilier du passage de lois prévoyant un investissement dans les infrastructures scolaires. De manière assez peu surprenante, les familles semblent donc valoriser monétairement la qualité des écoles mesurée en terme de résultats et d'investissement.

La mise en œuvre de procédures de choix scolaire centralisée a permis de mesurer de manière plus directe les préférences pour les attributs des écoles. En effet, dans le cadre de ces procédures, les parents doivent directement exprimer leurs choix en soumettant une liste de préférence au planificateur. Hasting et al. (2005) [84] utilisent la mise en place d'une telle procédure dans un comté aux États-Unis pour identifier les facteurs affectant les préférences. Ils trouvent que les parents valorisent la proximité, ainsi que les bons résultats scolaires. Cependant, ce sont avant tous les parents les plus riches et les plus éduqués qui valorisent l'excellence scolaire, et il existe une corrélation négative entre préférence pour la proximité et excellence académique.

Abdukadiroglu et al. (2014) [3] estiment un modèle de demande pour les écoles en utilisant les données du système scolaire de New-York, qui couvre un nombre d'élèves et d'école bien plus grand que celui étudié par Hasting et al. Ils trouvent de manière similaire que les les parents valorisent la proximité, ainsi que les résultats académiques et la composition ethnique de l'école. Là encore, les résultats tendent à varier par sous-groupe.

On peut se demander dans quelle mesure ces préférences pour les écoles sont rationnelles. En effet, comme on l'a vu, certains parents semblent peu sensibles à la qualité.

Hastings et Weinstein (2008) [85] répondent à cette question en menant une expérience contrôlée où des informations supplémentaires sur les écoles sont transmises à un groupe de parents choisis de manière aléatoire. Les informations portent en particulier sur les résultats académiques des écoles, et leurs capacités à faire progresser les élèves. Cette intervention a pour effet de changer de manière significative les choix des parents en faveur des écoles les plus performantes, l'effet étant particulièrement fort pour les familles les moins favorisées.

D'autres études montrent de manière similaire que les choix éducatifs ne sont pas forcément pleinement rationnels, et peuvent être modifiés de manière importante par des interventions changeant uniquement les informations disponibles pour les élèves et leurs familles. Ainsi, Bettinger et al. (2012) [23] trouvent un large effet à une intervention consistant à aider des individus à remplir leur demandes d'aide financière pour l'enseignement supérieur. Cette aide est suffisante pour augmenter de 8 points de pourcentage le taux de scolarisation dans l'enseignement supérieur ("*college*") alors même qu'elle n'apporte aucun bénéfice financier direct. La rationalité du choix dans un contexte éducatif peut donc être mise en question.

## 1.4 Les établissements peuvent-ils sélectionner correctement leurs élèves ?

Les parents d'élèves ne sont donc pas forcément rationnels. Qu'en est-il des offreurs d'éducation ? En effet, ceux-ci doivent également faire un choix parmi les élèves désireux d'intégrer leurs établissements. En particulier, ils doivent correctement évaluer les capacités scolaires de ces derniers.

Les études sur le sujet sont beaucoup moins nombreuses par rapport aux études sur les préférences des élèves et de leurs familles, et se sont surtout concentrées sur l'enseignement supérieur. Les décisions des établissements d'enseignement supérieur sont importante dans la mesure où ce niveau d'éducation n'est pas obligatoire, et représente donc un investissement qui n'est pas forcément rentable pour les étudiants. De plus, la probabilité d'échec n'est pas négligeable. Dans la mesure où les universités ont des informations privées sur la capacité à réussir des étudiants, leur capacité à les sélectionner de manière appropriée peut donc apporter un gain social non négligeable.

Cette question a été examinée principalement en lien avec les programmes de discrimination positive, qui permettent à des étudiants issus de minorités défavorisées d'intégrer des établissements prestigieux malgré des résultats universitaires comparativement moins bons. En effet, on peut se demander si cette admission préférentielle ne se fait pas au détriment de ces étudiants, qui sont par la suite comparativement plus exposés à l'échec que s'ils avaient choisi une université mieux adaptée à leur niveau. Par ailleurs, il est également possible que les bénéficiaires des programmes de discrimination positive

choisissent des spécialités académiques plus faciles, mais moins lucrative, suite à leur entrée dans une université ayant des standards académiques élevés, et ce afin d'éviter l'échec dans les majeurs les plus difficiles. Si cet effet est avéré, on peut alors considérer que la décision d'admission des universités inclue une part d'irrationalité, ou du moins d'inefficacité, puisqu'elle conduit à intégrer des individus à haut risque d'échec.

Arcidiacono et al. (2013) [14] examine un programme de discrimination positive mis en place à Duke University, aux Etats-Unis. Ils montrent que les étudiants admis à travers ce programme ont de moins bons résultats académiques, sont incapables d'améliorer leur classement au fil du temps et de rattraper leurs camarades, et sont beaucoup plus susceptibles de se réorienter vers des spécialités plus faciles, mais aussi moins lucrative, notamment au sein des humanités et des sciences sociales. Ces résultats sont confirmés par une étude de Arcidiacono et al. (2014) [13] qui étudie dans quelle mesure les étudiants issus de programmes de discrimination positive en Californie sont susceptibles de se réorienter. L'étude montre que ces étudiants, et plus généralement ceux qui sont les plus faibles par rapport à la moyenne lors de leur entrée à l'université, ont de plus grandes chances d'échouer ou de se réorienter vers des spécialités des sciences humaines et sociales.

Au delà de la mise en évidence de certains effets pervers de la discrimination positive, l'étude d'Arcidiacono et al. montre également que les universités américaines ont des pratiques de sélection qui peuvent laisser perplexe l'observateur européen ou asiatique, plus habitué à un système de concours. Ainsi, les étudiants ayant des capacités athlétiques, ou encore ceux ayant un parent membre du corps professoral se voient attribuer une priorité dans les procédures d'admission. Cet apparent favoritisme a cependant des résultats mitigés pour ces bénéficiaires, dans la mesure où ceux-ci ont plus de chance d'échouer ou de choisir une orientation moins prestigieuse ou lucrative.

La manière dont écoles et universités sélectionnent restent donc encore peu explorée, mais des résultats préliminaires tendent à montrer qu'elles n'est pas forcément la plus efficace possible en terme de réussite scolaires ou de préparation à la vie active. Le troisième chapitre de cette partie examine ainsi un programme de discrimination positive mis en place dans une grande école française. Il montre que certains étudiants sont admis alors même qu'ils ont un risque important d'échouer en raison de leurs caractéristiques observables. Par ailleurs, la procédure de sélection pour les autres étudiants n'est elle-même pas optimale dans la mesure où les étudiants issus de certaines voies d'entrée réussissent systématiquement mieux que d'autres.

## 1.5 L'allocation des élèves aux écoles dans un contexte de rareté

Quelque soit la rationalité des préférences, les satisfaire dans un contexte de rareté pose un problème supplémentaire. En effet, comment allouer les sièges d'une école trop demandée par rapport à sa taille ? Comment repérer les élèves qui ont la plus grande valorisation de cette école parmi tous les candidats ?

Depuis le début des années 2000, de nombreux travaux ont essayé de répondre à cette question, en s'appuyant sur la littérature sur les mécanisme d'appariement, initiée par Gale et Shapley (1962) [72]. Cette littérature est trop vaste pour être discutée dans le cadre de ce chapitre. Des revues de littérature très complètes ont été rédigées en anglais par Parag Pathak (2011) [134] et Abdulkadiroglu (2013) [2]. Cette littérature a été appliquées au cas français par Hiller et Tercieux (2014) [88].

La littérature sur le choix scolaire montre l'existence de compromis à trouver entre l'efficacité des mécanismes (aucun échange mutuellement avantageux possible entre élèves), leur stabilité (respect des priorités accordées à chaque élève), et leur résistance aux comportements stratégiques. En effet, il peut être intéressant de ne pas révéler ses vraies préférences dans le cas de certains mécanismes, et donc d'adopter un comportement stratégique. Il est alors possible d'essayer de voir quelles sont les familles les plus capables d'adopter un comportement stratégique, qui nécessite parfois des calculs complexes. Pathak et Sönmez (2008) [135] modélise par exemple la présence d'acteurs naïfs et d'acteurs stratégiques dans le système scolaire de Boston, et montrent comment les acteurs stratégiques sont capables d'améliorer leur situation au détriment des acteurs naïfs. He (2012) [87] estime un modèle de choix scolaire sur des données issues de la ville de Beijing en Chine, et met en évidence l'existence d'une hétérogénéité au sein des familles, qui sont plus ou moins prêtes à prendre des risques pour envoyer leurs enfants dans les meilleures écoles, et qui sont plus ou moins capable de mettre en œuvre des comportements stratégiques sophistiqués.

Deux mécanismes émergent donc de la littérature sur le choix scolaire. Le mécanisme d'acceptation différée (*deferred acceptance mechanism*) et le mécanisme de cycle d'échange (*top trading cycle*). Tandis que le premier respecte les priorités au prix d'une certaine inefficacité, le second ignore les priorités mais permet d'atteindre l'efficacité au sens de Pareto. Comme le montre Che et Tercieux (2013) [45], il s'agit également d'un mécanisme convergent asymptotiquement vers le maximum utilitariste, c'est à dire le mécanisme qui maximise la somme des utilités des individus. Ces deux mécanismes sont à l'épreuve des comportements stratégiques. Ils ne demandent pas de calculs sophistiqués aux familles.

On peut se demander dans quelle mesure un système centralisé est plus efficace qu'un système décentralisé où chaque élève doit faire un dossier de candidature pour chaque école. Abdulkadiroglu et al. (2014) [3] essaie de répondre explicitement à cette question

en comparant les allocations d'élève avant et après la mise en place d'un mécanisme centralisé. Il trouve que les élèves vont à des écoles plus lointaines, mais qui semblent mieux correspondre à leur vœux. De plus, moins d'élèves sont suffisamment insatisfait pour vouloir quitter le système public pour chercher une place dans une école privée. Par conséquent, un système centralisé semble supérieur à un système décentralisé.

Dès lors, un système public de choix semble supérieur à un système décentralisé se rapprochant plus d'un marché libre. Par ailleurs, un système centralisé permet de mettre en œuvre des politiques visant à réduire les inégalités entre élèves, en instaurant, par exemple, des quotas au sein de la population de chaque école, afin de promouvoir un plus grand mélange des étudiants. Enfin, comme le chapitre essaie de le montrer, il peut être nécessaire de manipuler les priorités données à certains élèves afin d'améliorer l'efficacité en terme utilitariste des algorithmes d'appariement.

Par conséquent, plutôt que du libre choix ou de la sectorisation, il semble plutôt qu'un choix organisé et régulé soit préférable.

## 1.6 Conclusion

Les partisans d'une dérégulation du marché scolaire mettent en avant la possibilité d'un meilleur appariement entre élèves et écoles comme étant un des possibles bénéfices d'une plus grande liberté de choix. Cependant, comme on l'a vu plus haut, en raison des contraintes de capacité des écoles, le choix scolaire est toujours bilatérale. Si les familles peuvent exprimer des vœux, elles n'ont aucune garantie de voir ces vœux exaucés. Dès lors, le libre choix décentralisé et la sectorisation ne sont que deux modalités possibles parmi de nombreux mécanismes d'allocation existant. En pratique, il semble que des mécanismes centralisés gérés par un agent public donnent de meilleurs résultats. On peut les considérer comme un marché, mais ce marché est de nature si artificiel que la distinction tranchée faite par beaucoup entre marché et État perd son sens. Par ailleurs, il peut être nécessaire de manipuler l'algorithme utilisé de manière importante afin d'améliorer le bien être social, comme essaiera de le montrer le second chapitre de cette partie.

De plus, on peut remettre en question la rationalité des décisions éducatives prises aussi bien par les familles que par les offreurs d'éducation. En effet, la plupart des acteurs du marché de l'éducation semble agir en présence d'information incomplète, et font preuve d'une rationalité limitée. Par conséquent, on ne peut qu'être sceptique quant à la capacité de l'exercice sans entrave du choix scolaire d'améliorer l'appariement entre élèves et écoles.





## Chapitre 2

# Matching with endogenously revealed preferences

## 2.1 Introduction

Since the pioneering paper of Abdulkadiroğlu and Sönmez (2003) [5], a series of papers have applied the tools of mechanism design to school matching. Why applied in nature, this literature has mostly focused on the theoretical properties of matching mechanisms. In particular, it has tried to find mechanisms which could respect axiomatic properties such that, for instance, the respect of existing priorities of students on school seats (stability), the inability for anyone to manipulate the mechanism (strategy proofness), or efficiency (no mutually beneficial exchange among matched students) (see for instance Pathak et al., 2008 [135], Kojima and Pathak, 2009 [108]). In its empirical part, it has studied how changing the use of matching mechanism could lead to different outcomes in term of respecting these criteria (see Abdulkadiroğlu et al., 2009 [4], Pathak and Sönmez, 2013 [136]). For instance, it has highlighted what could be the cost in term of efficiency of preferring mechanism guaranteeing stability rather than efficiency.

However, one of the important limit of this literature is that it has hardly been able to measure the welfare consequence of using one matching mechanism rather than another. In particular, there is no widely accepted criteria for comparing matchings (even though some papers did some proposals, Featherstone, 2011 [63]). Moreover, little is known on the welfare impact of the large change introduced by some school systems in their matching mechanisms.

As a result, there has recently been a surge of interest for models explicitly modeling the origin of the ranking submitted by individuals to the central planner. For instance, Che and Tercieux [45] (2013) analyzed a model where preference for schools are generated by a simple utility function with a common component and an idiosyncratic shock. They show how different mechanisms can converge or not toward the Utilitarian Upper Bound (that is the maximum utility which can be reached by maximizing the idiosyncratic component of preferences) when the size of the market grows. They also propose a new mechanism which is asymptotically stable and asymptotically efficient, that, is whose inefficiency and instability vanish when the market grows large.

While very interesting and innovative, the study of Che and Tercieux suffer from a limitation shared by a large part of the recent literature on school matching. It is based on asymptotic results which can apply only on large markets. For instance, the most discussed cases, theoretically and empirically, are those of the districts of New York, Boston and Chicago, which school dozen of thousands of students. While studying large markets has many advantages, particularly in showing how some inefficiencies can become fairly small when the market grows, it is at odd with the fact that most school markets are small, as documented in the first part of this paper. As a result, many asymptotic results may not hold in reality. Moreover, the literature on asymptotic properties of matching mechanisms hardly discusses the speed of convergence toward the desired property. It may be than a given mechanism takes a very long time to converge, and that some mechanisms having apparently the same asymptotic pro-

perties can sharply differ in term of speed of convergence.

A last limitation of the present literature on school matching is that it discusses only models where the revealed preferences for schools are exogenous, that is are generated by a process (for example, an underlying utility function) which does not depend on the matching itself. In particular, almost all the literature rests on the assumption that there is no interdependence across choices. When this assumption is relaxed, it is in a very abstract setting (Pycia, 2012 [137]), and by focusing mostly on the condition of the existence of an equilibrium. Therefore, in the best of my knowledge, no paper looks at the consequence of the presence of preference for peers on the efficiency of existing mechanisms.

The aim of this paper is therefore to propose a more realistic model of school matching, describing how preferences for schools are formed and including preferences for peers, and then to look at the welfare impact of the presence of preferences for peers, particularly in markets of realistic size. It also tries to see if the existing mechanism can be modified in order to improve the total utility of households when preferences for peers are present.

To do so, I propose a simple model where the preferences for schools are derived from a linear utility function which includes a common value component, a peer component, that is, how households value the characteristics of the students attending a school, and idiosyncratic preference shocks for schools which are randomly distributed. I make the crucial assumption that the shocks are correlated with some characteristics of the students which are judged desirable (or not) by all the students. I use either the Deferred Acceptance mechanism (DA) or the Top Trading Cycles mechanism (TTC) to match students to schools.

I then look for an equilibrium, where the expressed choices for schools and the resulting peer groups are mutually consistent. I show that a static equilibrium is unlikely to appear. I then re-frame the model into a dynamic setting where the current generation of students is making school choices based on past values of schools' characteristics. This generates a dynamic system which is stationary and oscillates around a steady state.

I show that if preferences are correlated, the steady state will exhibit school stratification : there will be a hierarchy of schools based on the desirability of their expected peer groups.

I then show by simulation the most important results of the paper : the presence of preference for peers and the presence of a correlation of idiosyncratic shocks with desired characteristics lead to a loss of welfare which does not disappear when the market grows reasonably large. This results hold regardless of the mechanism used, although the TTC mechanisms leads to a lower welfare loss. It may be that this loss vanishes asymptotically. However, it stays large even for the average school market size which is empirically observed.

It shows that preferences for peers can have important welfare consequences in matching problems. Even when the market grows large, they can still induce a non negligible welfare loss. Since as I show empirically, most markets are rather small, caring for the presence of preference for peers is relevant, and their absence in the literature on school matching is problematic.

This result is quite intuitive. The welfare gain that a mechanism can achieve compared to a random allocation depends on the correlation of the revealed preference for schools. In the extreme case where all students have the same ranking for schools, then no mechanism can do better in average than a random allocation of school seats to students. Preferences for peers can impact welfare because they increase the correlation between the ranking submitted by students to the social planner. In doing so, they reduce the informativeness of these rankings, since the planner can learn about the idiosyncratic preference for schools only thanks to the difference in ranking across individuals.

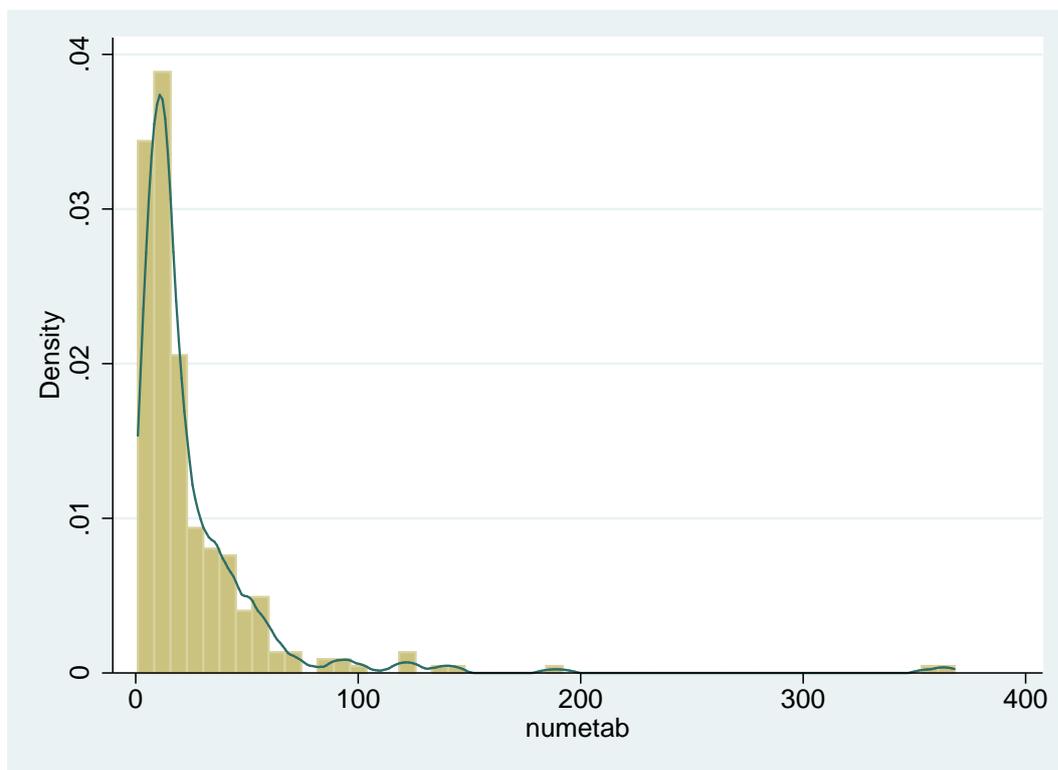
I propose a simple way to alleviate this problem. In order to reduce the correlation across choices, the planner can manipulate the random priorities which are drawn in order to break ties. By giving a higher priorities to the least desirable students, the planner can reduce the stratification across schools, and therefore reduce the correlation across preference. This improves the efficiency (in term of maximizing welfare) of the DA mechanism, but not of the TTC mechanism (probably because the latter already optimally trades priorities).

Beyond its contribution to the mechanism design literature, the paper has also implications in term of inequalities. Indeed, many policies have been implemented in order to reduce segregation, such as quota or busing. Very often, these policies proved to be controversial, and has been challenged by judicial decisions. This paper shows that simple modification to existing mechanisms can be implemented in order to limit inequalities across schools. Moreover, it shows that reducing between-school inequalities can improve overall welfare, which could made easier to gather supports for equalizing policies.

The paper proceeds as follow. In the second section, I present two stylized facts : most school markets are small, and parents care about the peers of their children. In the next section I present the model. The fourth section proposes some theoretical and numerical results. The last section concludes.

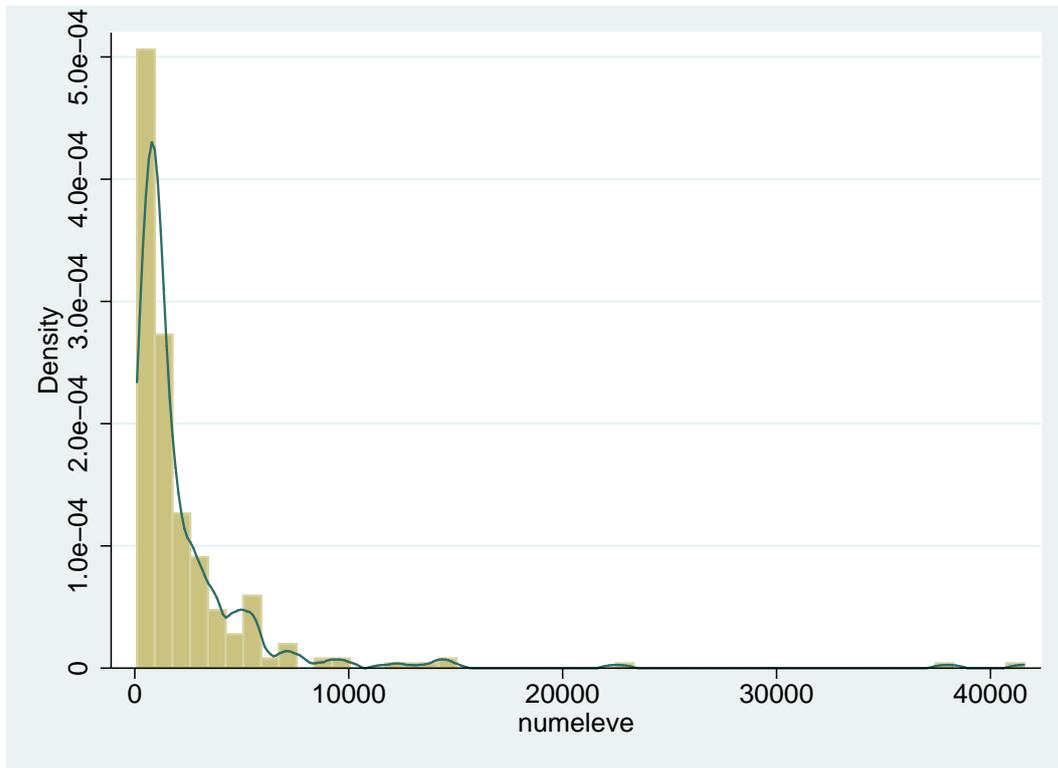
## 2.2 Some stylised facts : the distribution of the size of school markets and the existence of peer effects

### 2.2.1 The distribution of school market sizes



The graph above plots the distribution of junior high school market in France in 2010 by their size measured by the number of school. A school market is defined by using “*Zones d’emplois*”, or employment areas, which delimit different agglomerations in France. It corresponds to Metropolitan Statistical Areas in the United-States. School choice plan in France are in general organized at a higher level, the one of “*Académie*”, which corresponds to a region. However, in practice, students move very scarcely of cities when applying for schools. As a result, each school market can be consider as separate from the others, and the matching at the level of “*Académie*” can be thought of as an addition of separate matchings in each employment areas belonging to the “*Académie*”.

We can see easily that most school markets are small, with the median number of school equal to 15, the mean to 25, 75% of school markets below 30 schools, and 90% below 52. As a result, the large markets studied in the literature are an exception rather than the rule.



The graph above replicates the previous analysis, but school market size is now measured by the number of enrolled students. As before, the average market size is found to be small, equal to 2348 students, with a median at 1156, 75% of school market being below 2587 students, and 90% below 5120 students.

One could argue that even if there are not numerous, large school markets account for the majority of students due to their size. However, while they indeed school a large share of students, a significant share of students attend schools belonging to small markets. 14% of students attend a school in a market below the median, 32% a school in one of the 75% smallest markets, and 56% a school in one of the 90% smallest markets.

Moreover, using employment area is likely to overestimate the size of the school markets, since they can be geographically large. What's more, at the high school level, school markets become even smaller due to tracking, which separates students into virtually independent school systems, and because high schools tend to be much larger than junior high school (by a factor of two or three depending on regions).

### 2.2.2 Review of the evidence on the existence of peer effects

There is now a growing literature on parents and students preferences for school attributes. Hasting et al. (2008) [83] demonstrate that parents value schools having students with high scores, even though this preference is more pronounced for more educated and more affluent parents. They also show that minority parents face a trade-off between

choosing a school with high minority enrollment, and a school with better test scores (a direct example of the importance of preferences for peer attributes in matching problems). He (2012) [87] shows on Chinese data that parents value the average scores of school in a study on the strategic behavior of households in the Beijing school choice system. Similarly, Abdulkadiroğlu et al. (2014) [3] study the preferences for school attributes in New-York, and show that among other things, parents value school average test scores as well as demographic characteristics such as the ethnicity of the students attending the school.

In all these studies, parents value school attributes which depend on the choice of other parents, that is the average test score, and the share of students of various ethnic origins. Moreover, these attributes, and particularly test score, are shown to be comparatively important compared to other dimension of school quality.

Note that these preferences for peer attributes are not necessarily justified. Indeed, the literature on peer effects fails to find large and significant peer effects as far as cognitive skills are larger. The presence of peer effects of a quite large magnitude on behavioral outcomes such as crime or risk taking is however better documented. For this reason, the papers uses the expression “preferences for peer attributes” rather than “peer effects”.

## 2.3 A model of formation of school preferences

We know describe the model, which incorporates explicitly preference for peer attributes. A school market is made of the following elements :

**Schools :** There are  $S$  schools denoted by  $s \in S_s$ . Each school is characterized by its efficiency, or its commonly valued component,  $e_s$ . It has a fixed capacity, given by  $N/S$ , where  $N$  is the total number of students.  $S$  is typically quite small, most of the time less than a few dozen. We assume that schools have no preference on students.

**Students/households :** There are  $T$  generations of students, with  $N$  students denoted by  $i \in I_n$  in each generation. Each student is defined by his ability  $a_i$ . This characteristic is randomly distributed across students in the interval  $[\underline{a}, \bar{a}]$ , with a mean  $\mu_a$  and a variance  $\sigma_a$ . Each student has an idiosyncratic component affecting his valuation for each school, equal to  $\epsilon_{si}$ , a random variable. Students have a vector of priorities to school, with a students with a higher priority being preferred in case of a tie with another student having a lower priority. These priorities are denoted  $p_{is} \in P_s$ , and are also random variables. The priorities and the idiosyncratic components are correlated with ability, and potentially correlated with each other as well, with a variance - covariance matrix of the joint distribution between the  $\epsilon_s$ , the  $p_s$  and  $a$  given by  $\Omega$ .

Students derive utility by going to a given school  $s$ , given by the function  $U_{si}(e_s, \theta_{s(t-1)}, \dots, \theta_{s(t-t')}, \epsilon_{si})$ , with  $\theta_{s(t-1)}$  being average peer quality at time  $t-1$ , namely the average student ability in school  $s$ . As a result, households base their preferences on past values of the school characteristics. We assume that the function  $U$  is increasing and continuous in all its arguments, that's it is additive in its components, and that it is linear in the peer component, and that it is defined on  $R_+$ . An instance of such a function would be  $U_{si} = e_s^\alpha + \frac{1}{T} \sum_t \theta_{st} + \epsilon_{si}$ .

Note that since the peer components enter linearly into the utility function, the repartition of students in the school system is a null sum game as far as peers are concerned. If one student goes to a school with better peers, it means that another students will go to a worse school, since there is no optimal repartition of students across schools. This assumption is not very important by itself, but enables to clearly shows the impact of peer effects and preference for peer attributes on the maximization of the idiosyncratic component  $\epsilon_{si}$ .

**Matching process between schools and students :** Students order schools by preference, based on their utility function. Schools are assumed to not have any specific preference on students. A school  $s$  is preferred to a school  $s'$  if and only if :

$$U_{si}(e_s, \theta_{s(t-1)}, \dots, \theta_{s(t-t')}, \epsilon_{si}) > U_{s'i}(e_{s'}, \theta_{s'(t-1)}, \dots, \theta_{s'(t-t')}, \epsilon_{s'i}) \quad (2.1)$$

A matching mechanism  $\mu$  assigns students to schools, based on their submitted ranking and on their priorities. We consider two matching mechanism in the paper, the Top Trading Cycle algorithm (TTC) and the Deferred Acceptance mechanism (DA).

This simple setting departs in two major ways from the existing literature. First, revealed preferences for schools are now interdependent, and can change given the allocation of students to schools. As a result, a static equilibrium would exist only if there is an allocation of students to schools which generates the preferences needed to sustain the equilibrium.

Second, since this equilibrium is very unlikely to appear in practice, we depart from the usual static setting, and consider a dynamic model where students form their preference based on the past values of the peer groups  $\theta_s$ . As a result, the notion of equilibrium is replaced by the one of steady state, around which the system is going to oscillate. Indeed, each generation, the priorities and the idiosyncratic shocks are randomly drawn, which is going to make the matching between schools and students varies from years to years.

## 2.4 Numerical and theoretical results : the impact of preference for peers on welfare

Many papers on matching are concerned by the existence of an equilibrium. In practice, a static equilibrium is very unlikely to appear in our setting. Indeed, If the number of individuals and the number of schools are discrete, then, the set of all the possible peer groups generated by these two sets of individual and schools while respecting school capacity constraints is going to be discontinuous. As a result, suppose we draw a population of individuals with random shocks and preference, if we consider a mapping from the set of possible peer groups generated from this population to the resulting peer groups defined by a matching mechanism, then we have a mapping from a non-continuous set to another non-continuous set. This means that fixed point theorems such that Brouwer fixed point theorem or Banach fixed point theorem do not hold. In practice, simulation shows that equilibria are very unlikely to emerge as soon as the number of schools starts growing.

As a result, we do not try to establish the condition for the existence of an equilibrium, and we exploit the dynamic aspect of the setting to use the time series concept of asymptotic stationarity.

Define  $\bar{\theta}_{s,D} = \lim_{T \rightarrow +\infty} \frac{1}{T} \sum_t \theta_{st,D}$ , with  $\bar{\theta}_D$  the corresponding vector. This is the means of the time series of  $\theta_{st,D}$  defined by a given distribution  $D$  of  $e_s, a_i, \epsilon_i, p_i$ .

**Proposition 1 :**  $\theta_{st,D}$  is asymptotically second order stationary.

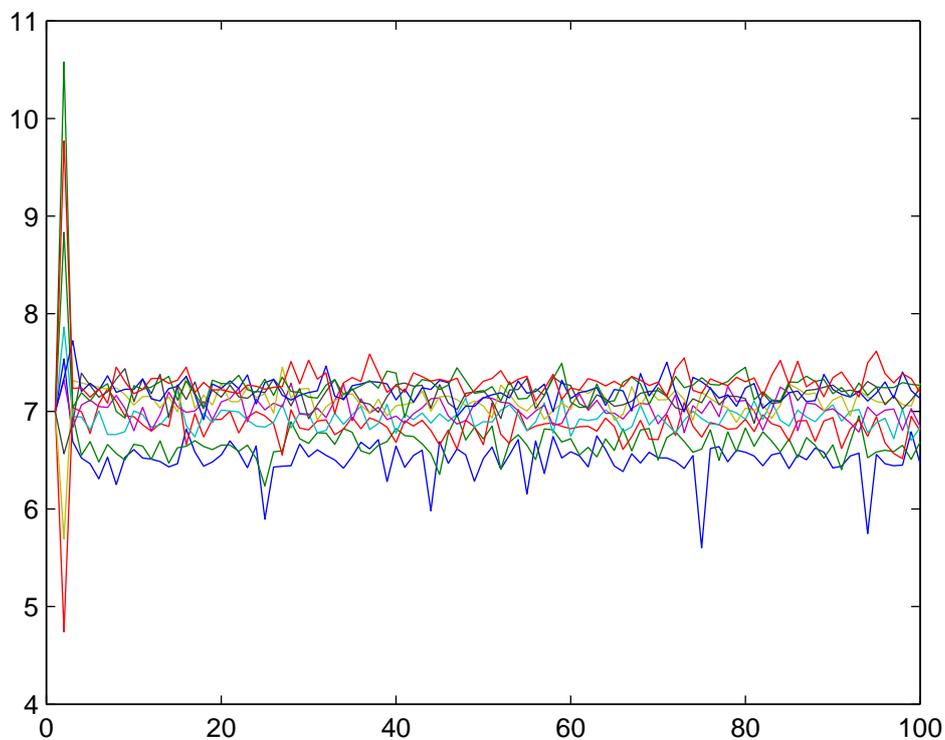
**Proof :** Any realized  $\theta_{st}$  is bounded above by  $\bar{a}$ , and below by  $\underline{a}$ . As a result, the mean on  $t$  of  $\theta_{st}$ ,  $\bar{\theta}_{s,D}$ , exists, and belongs to the interval  $[\underline{a}, \bar{a}]$ . This is enough to prove that  $\theta_{st,D}$  is first order asymptotically stationary. By the same reasoning, the variance of  $\theta_{st}$  exists, and belongs to the interval  $[0, (\frac{\bar{a} + \underline{a}}{4})^2]$ ,  $(\frac{\bar{a} + \underline{a}}{4})^2$  being the maximum possible variance, which occurs if the series alternate between  $\underline{a}$  and  $\bar{a}$ . This proves that  $\theta_{st,D}$  is asymptotically second order stationary.

The first proposition means that the system is going to evolve around a long-run trend, and will not diverge. The concept of asymptotic stationarity is much weaker than the one of a static equilibrium, however, it is also more realistic. Indeed, matching markets such as school markets are not perfectly stable over time and are affected by random shocks. Since households have limited information about the current demographic characteristics of the participants of the market, they will form expectation on the quality of school peer groups based on past values, which are easier to observe. As a result, the  $\theta_{st,D}$  will follow some sort of ARMA process, the stochastic component being provided by the randomness of the shocks, the abilities and the priorities.

Note that the result holds regardless of the process used by households to assess peer

group quality. We could, however, be willing to assume that families form rational expectations. In this case, we should specify a model used by families to extract information from the past such that there is no information left (for instance, an ARMA with the right number of lags). However, this model itself will have an impact on the matching. This complicates the problem because now we would have to find a predictive model consistent with the resulting matching (which would be equivalent to a rational expectation equilibrium). To avoid these complications, I stick to a behavioral interpretation of the model, where families use a simple rule of thumb to assess the future quality of the peer groups.

To illustrate the first proposition, I simulate the model during one hundred period, using the TTC mechanism to match students to schools. I show in the graph below how the behavior of the system looks like in such simulation. We can see that there is a stationary trend, but that there is also a major role of the random components.



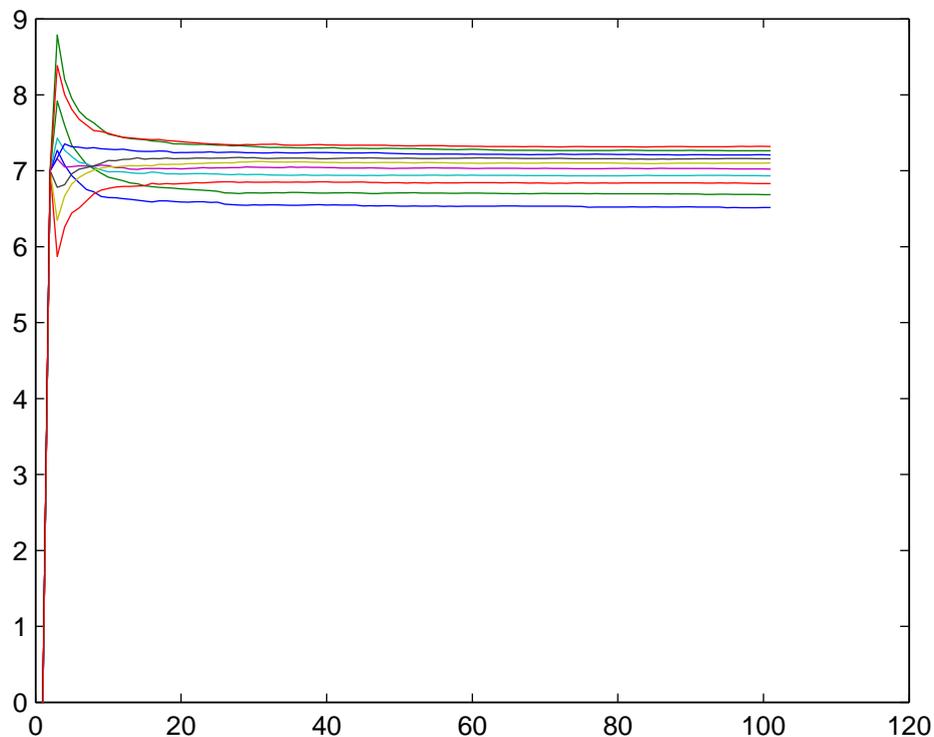
**Proposition 2 :** If  $corr(\epsilon_{is}, a_i) \neq 0$ , then peer groups are stratified, that is  $\bar{\theta}_{1,D} > \bar{\theta}_{2,D} > \dots > \bar{\theta}_{S,D}$

**Proof :** We proceed by building a proof by contradiction. Let's suppose that there is no stratification, that is  $\bar{\theta}_{1,D} = \bar{\theta}_{2,D} = \dots = \bar{\theta}_{S,D}$ . Then, peer groups will play no role in the choice of schools, which is going to be determined by  $\epsilon_{is}$  alone. Given that there is a positive correlation between shocks and priorities in one hand, and abilities

in the other, then the most able individuals are more likely to submit rankings which are correlated with each other. Therefore, since being admitted in a given school is positively linked with ranking high this school, both in the DA and the TTC, then, abler students will tend to gather in the same schools. But this contradicts our starting point that there is no stratification.

The second proposition means that in average, schools will be stratified by peers groups. This result is quite intuitive. Since preferences for schools are correlated with ability, high ability individuals will tend to gather in the same schools. However, this ranking can vary from year to year, since there is a large stochastic component in the determination of peer groups.

To illustrate the second proposition, I took the mean of all previous period at each point of the simulation. This shows how the means of school peer groups converge to a finite value, and how a clear stratification emerges across schools.



**Numerical result 1 :** The presence of preference for peers and of a positive correlation between idiosyncratic shocks and ability reduces total welfare compared either to no peer effects, or no correlation.

**Notes for a proof :** A possible proof could proceed as follow. First determine the utilitarian limit of a given mechanism. For the TTC, Che and Tercieux proves for instance that it converges toward the Utilitarian Upper Bound. Then prove that the

convergence process is monotonic. Then prove that the presence of preference for peers in a given finite market cause a loss in welfare.

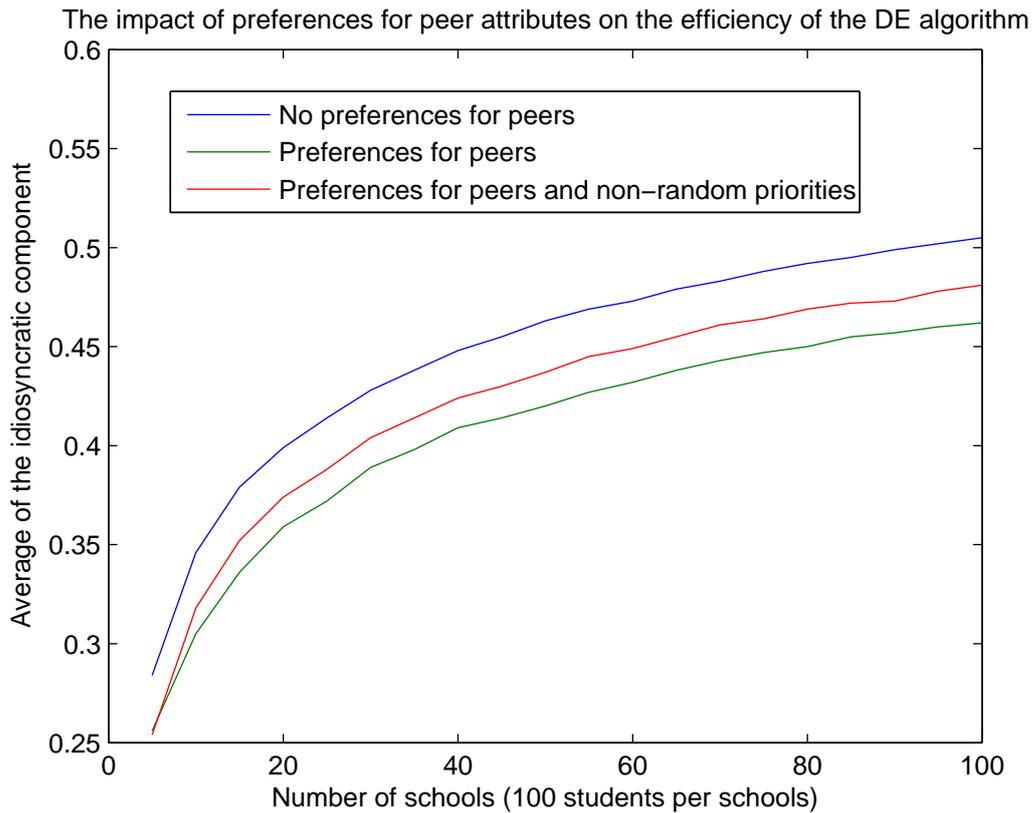
To do so, first prove that the presence of peer effect cause an increased correlation in the ranking submitted to the school. Then, show that the more correlated the ranking are, the lower the utility is. That's trivial in the extreme case (ie. comparing zero correlation to unit correlation), and quite intuitive in general. This could however be more difficult to prove generally.

**Numerical result 2 :** If we allocate the random priorities used to break ties in the TTC and the DA mechanisms such that we give the highest priorities to the least desirable students, then we can increase total welfare compared to a random allocation of priorities in tie-breaking in the case of the DA mechanism.

**Notes for a proof :** A possible proof could proceed as follow. First, we need to prove that allocating priorities preferentially to the least preferable student is going to decrease the stratification across schools. This is true because if a school has good peers, it will attract more students from all background, increasing competition for seats. However, since the least desirable students have higher priorities, they will be more likely to be admitted. This will drive down the quality of the peer group of the school, and thus stratification.

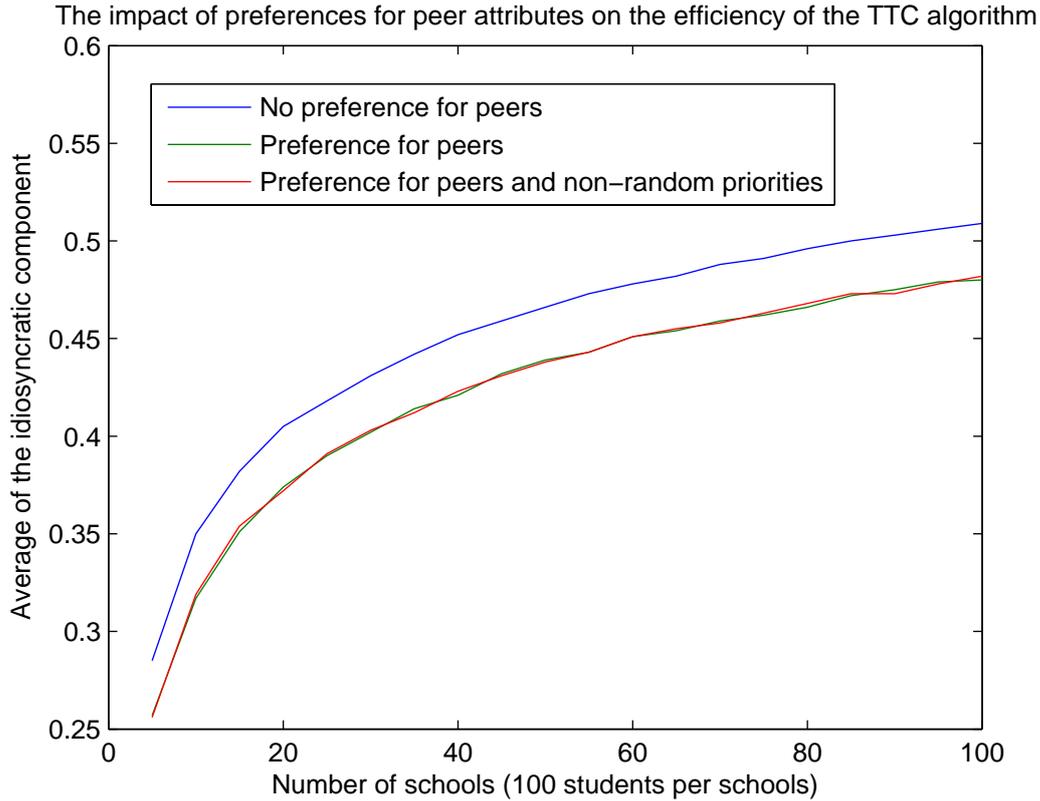
If we manage to prove that more stratification increase the correlation between submitted ranking, and that more correlated preferences hurt welfare (as we should do in the previous result to be proved), then we obtain the result.

To illustrate the two previous results, I show below the average utility in a market with growing size and using the Deferred Acceptance mechanism, with both random priorities (with and without preferences for peers), and preferential priorities for the least desirable (with preferences for peers). I compute the average value of the idiosyncratic component of the utility function across students for markets with 5 to 100 schools, each school having a size of 100 students. For each market size, I simulate the system during 50 periods, and then take the average on the last 30 periods.



The blue line shows the average individual utility when there is no peer effects, and the green line when there is preferences for peers, and random tie-breaking. The red line display the results when preferences for peers are present, and priorities are given to the least desirable students in tie breaking. In all case, the average utility grows with market size. However, the situation with preference for peers is always clearly worse. Giving priorities in tie-breaking to the least desirable students enables to improve the outcome, but still implies a loss of efficiency compared to the absence of preferences for peers.

I performed below the same analysis for the TTC mechanism.



The results are quite different from the simulation with the DA mechanism. As for the DA, the welfare without preferences for peers is higher than with preferences for peers. However, the difference between the case with preferences for peers and with and without priorities given to the least desirable is very weak, with no mechanism unambiguously dominating the other. This could be due to the fact that the TTC mechanism is already trading optimally priorities given to individuals, and that as a result, manipulating priorities does not bring any improvement. As a result, the TTC seems better at dealing with the inefficiencies generated by the presence of peer effects. However, it is unable to completely suppress these inefficiencies, even for markets with large size. As a result, one could be skeptical concerning its ability to converge asymptotically to the utilitarian upper bound .

The parameter used in both simulation are reported in the appendix.

## 2.5 conclusion

In this paper, I investigated the welfare impact of the presence of interdependent preferences (i.e. preferences for peers) on markets of small to average size. This is motivated by the fact that empirically, most school markets are small, and that households have been consistently found to value the presence in the school of their children of high achieving students, as well as student with characteristics similar to the one of their children.

While these two features are hardly discussed in most of the literature, they are shown to be important for welfare. Simulations show that the presence of preferences for peers lead to a loss of welfare both with the DA and the TTC mechanisms. This loss does not seem to vanish when the market grows large, and it may be that it is still present asymptotically.

The intuition for this result is simple : preferences for peers increase the correlation among the school rankings submitted by the parents to the social planner. This prevents the planner to gather information on the idiosyncratic preferences of the parents for the schools.

I show how manipulating the random priorities used to break ties can alleviate the inefficiencies in the case of the DA mechanism by giving priority to the least desirable, and thus reducing the level of stratification across schools at the steady state. In doing so, I bring an important argument in the debate about the effect of segregation on welfare. Most of the discussion about segregation focuses on the inequalities it induces in the presence of peer effects. However, another problem generated by segregation is that it reduces the efficiency of market mechanisms and prevent students from being matched to their most preferred schools, based on their idiosyncratic preference only. Indeed, segregation creates a vertical differentiation across schools which leads parents to ignore their horizontal differentiation when choosing a school.

The paper also shows that simulations are important to assess the effects of mechanisms on markets with realistic sizes. Different types of markets may need different types of mechanisms, showing the importance of exploring further the convergence properties of matching mechanisms.

## 2.6 Appendix

**1. Parameters of the simulation in part 4 :** For the simulation, I use a simple linear utility function :

$$U_{ist} = e_s + \phi \left( \frac{1}{t-1} \sum_{t=1}^{t-1} \theta_{st} \right) + \epsilon_{is} \quad (2.2)$$

The household infers the quality of the students attending the school by the mean of the past values.

The ability distribution is normal, with a mean  $\mu_a$  and a variance  $\sigma_a$

The shock  $\epsilon_s$  are normally distributed with a variance  $\sigma_\epsilon$ , common to all the shock, a covariance  $\sigma_{\epsilon\epsilon'}$  across shocks for two different schools, and a covariance  $\sigma_{\epsilon a}$  across shocks and abilities.

There is a variance / covariance matrix between  $a$  and the  $\epsilon$ , denoted  $\Omega$

I fix the parameters as follow :

- $e_s = 1$  for all  $s$
- $\phi = 0.6$
- $\mu_a = 7$
- $\sigma_a = 1$
- $\sigma_{\epsilon_{is}} = 0.15$  for any  $s \in S$
- $\sigma_{\epsilon_{\ell'}} = 0$
- $\sigma_{\epsilon_{i,sa}} = -0.5 + s/S$  with  $S$  the number of schools





## Chapitre 3

Mismatch or convergence ?

Affirmative action and students'  
selection in a French elite university

## 3.1 Introduction

Affirmative action in higher education, that is the preferential treatment given to members of historically deprived groups during the admission stage at selective universities, is a popular way to improve the outcome of the members of under-represented minorities. Indeed, top higher education institutions are often the pathway to the most preferred positions in society. Moreover, a seat in such universities can also be seen as a valuable investment good, since students of top universities can profit from better peers and better teaching conditions. Moreover, occurring soon in life, such advantage can change the entire life course of an individual, and help to close the gap between the targeted minority and the majority.

However, while widespread, such affirmative action programs have met fierce resistance. Recently, several judgments in the US have banned the implementation of affirmative action programs in higher education system, in states such as California or Michigan. Most of the arguments for or against affirmative action programs are deontological in nature, and beyond the scope of economics or empirical research. However, two empirical arguments have been discussed in the debate about affirmative action.

First, supporters of affirmative action have argued that the assessment of candidates to university is mostly based on their contemporary achievement, and does not take into account their potential. However, students coming from disadvantaged ethnic or social groups have low achievement because of their current economic and social situation. As a result, in a more favorable environment such as the one of a prestigious university, these students should be able to catch up and to reveal their full potential. Therefore, affirmative action is not only a way to compensate the members of disadvantaged groups, but is also a way to recognize their underlying abilities. This argument is empirical in nature because it relies on the fact that students admitted through affirmative action programs are able to catch up during the course of their studies with regularly admitted students.

In the opposite, the critics of affirmative action argue that it leads to a mismatch between the desired study plans of the beneficiaries and their abilities. Being unable to fulfill the requirements of the selective institutions where they have been admitted, and missing the private information that universities have on the difficulties of the curriculum to do an enlightened choice (Arcidiacono et al. 2011) [12], affirmative action beneficiaries are more likely than other students to either fail or to switch to an easier and often much less lucrative major. As a result, affirmative action is a mixed blessing for those who benefit from it, and could in fact worsen their academic and professional outcome. Again, this is an empirical argument.

More generally, the arguments above lead to question the current design of affirmative programs. Indeed, many of these programs are based exclusively on ethnic belonging. As a result, they can lead to give an advantage to students who, while coming from a deprived group, are themselves very well-off. Moreover, they do not try to specifically

target applicants from deprived groups who are likely to catch up, for instance because they come from a negative environment. As a result, one can wonder what would the best policy to preferentially admit students with weak credentials, but strong potential to improve.

The aim of this article is thus to bring empirical arguments to this debate by studying a specific affirmative action program implemented in a French elite university. This program is interesting because it differs from traditional program who focus on ethnic group belonging, and rather target high schools in deprived area, whose students come from a variety of ethnic background, but are almost all belonging to poorer than average families. In practice, this includes many students from immigrant families, thus having a similar effect as a traditional affirmative action program. As a result, it differs from the practice of many top American universities who admit average applicants coming from middle class background due to race preferences in admission. Given that these applicants do not come from a deprived background which could have hindered their past academic performance, they are less likely to be able to catch up. As a result, studying the program implemented by this French university give us the opportunity to assess what is the best design of an affirmative action program.

Can a program based mostly on social background do better than a program only based on ethnic affiliation? Are affirmative action students recruited through this procedure more likely to catch up compared to the experience of US universities? Does this university successfully target the students who are the most likely to improve?

To answer these questions, this article first examines if catching up is occurring in this elite higher education institution. The main problem in assessing difference in achievement across group is that the selection into courses could be endogenous (Arcidiacono et al. 2012) [14]. For instance, affirmative action students could choose easier courses, or courses attended by relatively weak classmates, making it easier to achieve a good ranking. Moreover, the variance of grades could evolve through time, generating a mechanical convergence in grades across groups (though not in ranking). To overcome the first problem, I use the fact that the students in the university we study follow a curriculum core in their first year, and are randomly assigned to tutorials. When we measure again their achievement, in their fourth year of study, they have to attend programs with large common cores, which reduce the sorting problem within the program. As for selection into program, I can deal with it by measuring program fixed effects while controlling for first year achievement.

I show that there is a very large gap in the first year across groups, and a much higher rate of failure among affirmative action students. However, among the students who manage to reach in time their fourth year of study, the gap is much smaller, and there is some convergence with their peers. This is true both by measuring the gap in absolute value, and by computing ranks. As a result, it shows that if affirmative action students are in general weaker than their peers, there is much heterogeneity among them, and

that those who manage to progress until the end of the program can close part of the gap with their peers, with about one half of the gap in test scores and one third in rank erased after four years of study.

I then look at the determinant of academic achievement in first year and of convergence among affirmative action students. I find that there are some strong predictors of failure, in particular the major chosen in high school, with humanities major being much more likely to fail. I do not find strong predictors of convergence. As a result, it seems, surprisingly, that the university is ignoring valuable information during the selection process.

I also look at major choice of students to see if a mismatch is occurring. There is no strong choice pattern of major across affirmative action students, who are only less likely to choose more academic and less applied major, probably because they value more marketable skills. As a result, if any, there is a positive sorting of affirmative action students toward the most lucrative majors.

As a result, this paper sheds a new light on affirmative action program. It confirms the previous findings that affirmative action beneficiaries have lower academic performance than their peers. However, it also shows that a program emphasizing selection of beneficiaries based on their social background rather than uniquely on their ethnic origin is more likely to produce convergence, and reduce the risk of mismatch. It also demonstrates that universities do not always use all their information to select students.

The paper is related to the literature on mismatch in affirmative action. A first stream in this literature, using national surveys, shows how students, and in particular black students, who are attending schools while having comparatively low academic credentials are more likely to fail to graduate, or to graduate with delay (Loury and Garman 1995 [121], Bound et al. 2010 [30], Light and Strayer 2000 [118], 2002 [119]). Another part of the literature looks at how affirmative action bans affect minority enrollment and graduation. In general, these studies find no effect of these bans on graduation, which somehow contradict the mismatch hypothesis (Backes, 2012 [18], Hinrich, 2012 [89]). A last stream of papers use micro-data at the university level, which enables to examine not only graduation but also the sorting into majors. For instance, a paper by Arcidiacono et al. (2012) [14] documents how black students at Duke University, who massively benefited from affirmative action during the admission process, fail to catch up with regularly admitted students, and frequently end up switching from science majors toward less challenging but also low return humanities and social science degrees. These findings are confirmed on a larger scale by Arcidiacono et al. who document the pattern of choice of major in Californian universities (Arcidiacono, 2013) [13]. The literature using university level data also includes the literature on mismatch in law school (for instance, Sander (2004), [153], Antonovics and Sanders (2013) [9], Ayres and Brook (2005) [17], William (2013) [158] and Rothstein and Goon (2008) [150]) which tries to measure the impact of affirmative action on academic results, graduation and

success at the bar exam, but which deal with different econometric problems, related to the fact that some recipient of the affirmative action programs benefit of it at the extensive margin, which make difficult to build a valid counter-factual. It shows that if these students are more likely to be at the bottom of the distribution of score, they would anyway have had no chance to be admitted at all in any law school without affirmative action. As a result, it is unlikely that they are hurt by affirmative action, which indeed give them an opportunity to access prestigious but selective legal jobs.

The paper proceeds as follow. Section II presents the data and the characteristics of the affirmative action program. Section III explore the evolution of the gap in academic results and the choice of major across groups, and examine the existence of a convergence process. Section IV looks at the determinants of academic success and convergence among students, and tries to see if there are some predictors of failure among them. Section V concludes.

## 3.2 Data and institutional environment

### 3.2.1 The affirmative action program

The university that we study is specialized in social sciences. It is very famous in France for training most of the high rank officials of the public administration, and many politicians. For instance, the last four French presidents graduated from this university, as well as many of the members of the parliament. The school is very selective and belongs to the “*grandes écoles*” system, which gathers the French top students. Thus, about 70% of the regular students passed the national high school graduation exam with highest honor (“*mention très bien*”), a performance achieved by less than 2% of their age group. During the time span of the study, the school was admitting between 500 and 1000 students each year. Among French elite university, it is among the only one recruiting immediately after high schools, which makes easier to implement an affirmative action program.

The program was implemented in 2001, in response to frequent criticisms about the lack of social openness of the university, and the detrimental effect it had on the composition of the French political elite, deemed too homogeneous. Indeed, the existing recruiting process consisted in two different paths : either an entrance examination or an application reserved only to students having passed the national high school graduation exam with highest honor. In practice, while anonymous and purely academic, the procedure led to a very large over-representation of students with wealthy background. This raised the issue of a “reproduction” of the existing elite, whose members were able to transmit their positions to their children, while students from middle and working class background had very little chance to enter prestigious universities.

Starting from 2001, a reform of the admission procedure was implemented in order to make the student body more diverse. The aim of the reform was slightly ambiguous, because it is strictly forbidden in France to use ethnic origin for running a public policy. Actually, even gathering data on ethnic origin is forbidden, since the French constitution explicitly rejects the idea of race. In general, race, and by extension, ethnic origin is seen as an unscientific concept mostly used for discriminatory purpose. As a result, the target group of the reform was defined by its social origin. The aim was to increase the number of students from the working class.

To do so, rather than explicitly favoring these candidates during the normal admission procedure, the university designed a special recruitment procedure which was opened only to students coming from a selected group of high schools, all belonging to “special education” (*“éducation prioritaire”*), that is receiving some additional subsidies from the state because of their high share of poor, immigrant, and low-achieving students. The program started with only a small number of students and high schools in the beginning (only seven high schools and less than thirty students) and grew until including several dozen of high schools, and admitting more than one hundred students yearly a decade later. As a result, it differs strongly from traditional affirmative action program, because it does not target a very specific ethnic or social group. However, in practice, the majority of the beneficiaries were from immigrant background, which means that the program replicated the effect of traditional affirmative action programs.

The admission procedure was organized in several steps. First, students had to enroll into a preparation program during the last year of high school, with the agreement of their teachers. This program provided information on the procedure to the students, and a preparation to higher education. The students willing to candidate had to prepare a report on an actuality issue, based on personal research, and then present it in front of a committee. Based on their performance and on their previous academic records, the committee then decided to admit them or not in the university. The procedure was rather selective, with between 15 to 20% only of the applicants who were admitted. The committee focused not only on academic results, but also on the attitude of the students and their motivation. While the program included only a small number of high schools and students in the beginning, it grew quite fast until the affirmative action beneficiaries constituted between 10 to 15% of the students enrolled at the bachelor level.

### **3.2.2 Data**

We obtained data on two cohorts of students who were admitted at the bachelor level in 2007-2008 and 2009-2010. We can observe some characteristics such as if they passed high school graduation exams with highest honor or not, an indicator of their social origin, how they were admitted in the university, their geographic origin, their gender,

and their nationality. We present some descriptive statistics below.

VARIABLES	(1) Admission through regular procedures mean	(2) Admission through the AA procedure mean
results 1st year	13.05	10.68
HS humanities track	0.152	0.137
HS Scientific track	0.407	0.251
HS social sciences track	0.428	0.592
Other track in HS	0.0101	0.0190
Foreign HS degree	0.00289	0
Overseas HS degree	0.0145	0.114
HS degree obtained in a Foreign country	0.0116	0
HS degree obtained in Paris region	0.444	0.649
Highest honor at HS graduation exam	0.745	0.209
Female	0.563	0.630
Foreigner	0.0188	0.0474
Dual nationality	0.0419	0.123
<hr/>		
<b>Father occupation :</b>		
Farmer	0.00795	0.00474
self employed or running a company	0.0766	0.0474
Intermediate job	0.0918	0.137
Employee	0.0376	0.218
Worker	0.00578	0.142
Retired	0.0441	0.109
Inactive	0.0159	0.0853
Missing	0.0217	0.0616
Professional	0.699	0.194
<hr/>		
<b>Mother occupation :</b>		
Farmer	0.00361	0.00474
self employed or running a company	0.0397	0.0190
Intermediate job	0.133	0.109
Employee	0.0679	0.199
Worker	0.00434	0.0853
Retired	0.0361	0.0284
Inactive	0.111	0.166
Missing	0.108	0.261
Professional	0.496	0.123
Observations	1,384	211

We can see that affirmative action students come from a much less wealthy and educated background than their peers. In particular, while 70% of the regular students have a father who is professional, this is the case for only 20% of the affirmative action beneficiaries. They are also academically weaker, which much less of them having obtained highest honors while graduating (20.9% against 74.5%). I can follow these students during the course of their studies. We had access to their results at the end of the first year (mean GPA). I can then know how they spend their third year, since studying or doing an internship abroad is compulsory for all students. I then have their results at the end of their fourth year in the university (corresponding to the first year of a master's degree in France), as well as their choice of master. At a first glance, we can see that affirmative action recipients have worse results than the others, but show improvement over time.

### **3.3 Do affirmative action students catch up ?**

#### **3.3.1 Identification strategy**

I first try to determine how affirmative action students perform academically, and to see if they can catch up with their peers. The main econometric problem is linked to the non random selection of courses and majors. Some students could systematically choose easier courses, allowing them to obtain better marks. Moreover, if students sort across courses based on difficulties and abilities, the relative ranks of students would be an incorrect measure of achievement, because, for instance, weaker students could take the same courses, and then more easily achieve a good ranking because they are not competing with better peers. To deal with this problem, Arcidiacono et al. (2012) [14] compute student fixed effects while controlling for course fixed effects and adjusting for truncation due to grade inflation. They can then rank students based on their fixed effect, and thus obtain an accurate and unbiased measure of convergence.

In this paper, I do not need to adjust for truncation, because no students are able to obtain the maximum mark (grade inflation is non existing in France). I cannot control for course difficulty, nor compute student fixed effects, because we only have an aggregated measure of academic achievement, that is the grade average in first and fourth year. However, I can take advantage of the particular structure of the curriculum in order to eliminate most of the bias associated with course selection.

Indeed, the first year of studies is constituted by a curriculum core where all students have to take the same courses, with no choice left (with the exception of a few special programs where students take additional courses, for which I can control for). The only element of differentiation are the tutorials that students have to take, and which determine two third of their final grade in a given course. However, this choice can be considered as being random. Indeed, students cannot enroll to a tutorial independently,

but have to take together a pack of three tutorials (which have the nickname “triplette”) at the same time each semester, covering all the subject being studied. The tutorials are grouped in pack mostly in order to strengthen the link between students, and also in order to balance schedule. When this system was implemented and until several years later, it was impossible to have access simultaneously to the schedule of all tutorial in a same pack. As a result, comparing the value of each pack in term of schedule was very difficult. Moreover, students do not have information about teachers’ quality and grading policy. Indeed, there is a large turnover of instructors, and the enrollment is done before the beginning of the academic year. As a results, students have little possibility to acquire information about teachers by interacting with past first year students.

What’s more, the enrollment is done on line, and all students have to enroll at the same time during a very short time span (few minutes in practice). Since each vacant slot is fulfilled very quickly (sometimes in few seconds), it is difficult for students to coordinate in order to be enrolled in the same pack. Indeed, the data show no strong apparent sorting across groups based on observables (see Appendix I). We can only notice that affirmative action student tend to avoid some groups, probably because of their schedule. Indeed many of them come from northern Paris, which makes difficult to attend courses very early or very late. We have also one group with a very large share of affirmative action students, which may be due to coordination. However, such a coordination across students could have an effect only if there is peer effects. However, as we will see later, we do not find any peer effects, in line with most of the literature on peer effects in higher education (see for instance Sacerdote, 2001 [152], Zimmerman, 2003 [159], Stinebrickner et al., 2006 [157], Lyle, 2007 [122], and Foster, 2006 [68]). As a result, student sorting into group is very unlikely to create any bias, and it is possible to unbiasedly estimate the academic gap across groups by regressing student scores on students’ admission procedure.

At the master level, however, there is more room for sorting to occur. Indeed, students have to choose a master specialization, and inside each master, there is only a part of the curriculum which belongs to a core compulsory for all students. However, this core is typically large, including 60 to 80% of the academic credits of the two semesters of the first year of master, which mitigates the sorting problem.

In order to estimate consistently the gap between each group of students, I first use master fixed effects in the regression of test scores on students’ admission procedure. This eliminates the influence of the difficulty of the master degree on scores. However, it does not suppress the effect of the sorting of students into masters based on ability. Thus, the estimates of the master fixed effects could be biased (for instance, if weak students selected on easy masters, the fixed effects estimates of these masters would be biased downward). In order to control for this I estimate a specification where I control for both master and first year test scores. Thus, I can obtain unbiased estimates of the master fixed effects. Then, I can estimate a constrained regression where I constrain

the master fixed effects to be equal to their estimates in the specification with first year test scores. If I find large differences in the estimate of the effect of the admission procedures, it shows that a sorting is occurring.

### 3.3.2 The academic achievement of affirmative action students

I present some results below. First, I compute the effect of the admission procedures on the results in first year on the entire sample.

As we can see, students admitted through the affirmative action procedure perform much worse than the others. The magnitude of the difference is large, and equal to 102% of a standard deviation. Controlling for some demographic and academic characteristics such as gender, if the students graduated with highest honor, and the place where he or she took the end of high school exam explain about one third of the gap. However, controlling for the socioeconomic status of the two parents, and with peer characteristics does not change very much the estimates.

VARIABLES	(1) Result 1st year	(2) Result 1st year	(3) Result 1st year	(4) Result 1st year
Admission : Highest Honors	0.737*** (0.101)	0.274** (0.110)	0.305** (0.116)	0.306** (0.123)
Admission : Math track	-0.275*** (0.0751)	-0.612*** (0.0832)	-0.628*** (0.0870)	-0.245 (0.318)
Admission : humanities track	-0.306 (0.424)	-0.132 (0.460)	-0.126 (0.425)	-0.00512 (0.400)
Admission : Science track	0.391*** (0.144)	-0.219 (0.189)	-0.220 (0.185)	-0.981** (0.375)
Admission : Affirmative action	-2.196*** (0.242)	-1.469*** (0.197)	-1.287*** (0.212)	-1.288*** (0.211)
Constant	12.75*** (0.0908)	12.09*** (0.144)	12.24*** (0.160)	13.45*** (1.218)
Characteristics	NO	YES	YES	YES
Father occupation	NO	NO	YES	YES
Mother occupation	NO	NO	YES	YES
Peer Characteristics	NO	NO	NO	YES
Observations	1,614	1,595	1,595	1,595
R-squared	0.212	0.316	0.334	0.341

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

We can see that other tracks of admission can also have lower results in first year. In particular, all the students admitted by partner institutions into dual programs (Math, humanities, and science track) perform less well, even though the results are not always significant. However, the magnitude of the gap is much lower compared to affirmative action students.

Not only the grades of affirmative action students are lower, but they are much more likely to fail to pass the exams and to be unable to pursue in second year, which implies that they have to repeat or drop out, as it is shown below in a logistic regression. Since course allocation is random, this effect can entirely be attributed to a lower academic level, a lower motivation, or a lower preparation of these students. The results are robust to the addition of the average quality of the peers, class size, gender, major in high school, social origin, nationality and tuition fees. For each specification, we also add year fixed effect to control for events affecting a single cohort. We can also notice that there are differences across admission procedures. In particular, students entered through exams perform less well than students entered through the application procedure.

VARIABLES	(1) fail "premier cycle"	(2) fail "premier cycle"	(3) fail 1st year	(4) fail 1st year	(5) conditional passing	(6) conditional passing
Admission : Highest Honors	-0.304 (0.320)	0.563 (0.512)	-0.605* (0.333)	0.273 (0.485)	-0.389** (0.176)	-0.0672 (0.244)
Admission : Math track					0.0850 (1.064)	2.258* (1.322)
Admission : Science track	-1.020 (1.028)	2.850 (8.979)			-2.354** (1.015)	0.563 (2.773)
Admission : Affirmative action	1.729*** (0.271)	1.168** (0.458)	2.016*** (0.252)	1.253*** (0.410)	0.426** (0.210)	0.386 (0.290)
Admission : humanities track	0.984 (0.781)	0.812 (1.052)	0.781 (0.777)	0.562 (1.034)		
Constant	-3.557*** (0.267)	-8.727 (6.472)	-3.660*** (0.269)	-5.105 (5.876)	-1.407*** (0.134)	2.390 (2.925)
Father occupation	NO	YES	NO	YES	NO	YES
Mother occupation	NO	YES	NO	YES	NO	YES
Peers	NO	YES	NO	YES	NO	YES
Characteristics	NO	YES	NO	YES	NO	YES
Observations	1,594	1,544	1,524	1,458	1,593	1,566

Standard errors in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

However, I cannot use this sample to measure convergence, because I cannot observe the fourth year results of the students who have dropped out or repeated a year because of a failure to pass exams during the first year. As a result, I estimate the same specification on a sample restricted to the students able to pursue until the fourth year of study. As one could expect, the gap between groups, while still large, is much lower, at about 80% of a standard deviation. The difference in absolute value is even stronger, because removing students who fail change in a sizable way the standard deviation of the results.

VARIABLES	(1)	(2)	(3)	(4)
	Result 1st year	Result 1st year	Result 1st year	Result 1st year
Admission : Highest Honors	0.573*** (0.0788)	0.270*** (0.0923)	0.284*** (0.0930)	0.269*** (0.0975)
Admission : Math track	-0.822*** (0.0590)	-0.957*** (0.0593)	-0.953*** (0.0733)	-0.963*** (0.208)
Admission : humanities track	-0.157 (0.308)	-0.124 (0.310)	-0.102 (0.298)	-0.0894 (0.233)
Admission : Science track	-0.00857 (0.133)	-0.358** (0.142)	-0.358** (0.143)	-0.745* (0.390)
Admission : Affirmative action	-1.068*** (0.107)	-0.753*** (0.109)	-0.751*** (0.147)	-0.792*** (0.144)
Constant	12.99*** (0.0583)	12.38*** (0.108)	12.42*** (0.119)	12.72*** (0.995)
Characteristics	NO	YES	YES	YES
Father occupation	NO	NO	YES	YES
Mother occupation	NO	NO	YES	YES
Peers	NO	NO	NO	YES
Observations	1,317	1,311	1,311	1,311
R-squared	0.178	0.273	0.284	0.298

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

We then estimate the impact of the admission procedure on the results in fourth year. We use several specifications with admission procedures fixed effects only, master fixed effects, and a constrained regression with master fixed effects previously estimated while controlling for test scores. The results for the impact of having been recruited through the affirmation procedure do not differ a lot across specifications, showing that sorting across masters is limited.

We find again a sizable gap between affirmative action students and their peers. However, when measured in standard deviation, it has almost been divided by two, at about

40%. As a result, it shows that there is a sizable convergence conditional on being able to pass the exams of the first year.

VARIABLES	(1) Result 4th year	(2) Result 4th year	(3) Result 4th year
Admission : Highest Honors	0.325*** (0.0962)	0.135 (0.143)	0.152 (0.150)
Admission : Math track	0.110 (0.499)	0.0170 (0.349)	0.0290 (0.294)
Admission : humanities track	0.438 (0.350)	0.369 (0.351)	0.371 (0.361)
Admission : Science track	0.683*** (0.174)	0.383* (0.185)	0.377* (0.195)
Admission : Affirmative action	-0.631*** (0.147)	-0.423** (0.146)	-0.355*** (0.116)
Constant	12.93*** (0.0394)	12.36*** (0.146)	12.40*** (0.164)
Masters dummies	YES	YES	YES
Constrained	NO	NO	NO
Masters dummies			
Characteristics	NO	YES	YES
Father occupation	NO	NO	YES
Mother occupation	NO	NO	YES
Observations	1,316	1,310	1,310
R-squared	0.094	0.129	0.153

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

VARIABLES	... Continued		
	(4)	(5)	(6)
	Result 4th year	Result 4th year	Result 4th year
Admission : Highest Honors	0.338*** (0.0971)	0.139 (0.143)	0.157 (0.149)
Admission : Math track	0.220 (0.385)	0.0912 (0.268)	0.0698 (0.248)
Admission : humanities track	0.364 (0.330)	0.305 (0.330)	0.322 (0.352)
Admission : Science track	0.754*** (0.225)	0.409* (0.209)	0.396* (0.207)
Admission : Affirmative action	-0.667*** (0.129)	-0.443*** (0.132)	-0.374*** (0.108)
Constant	12.86*** (0.0482)	12.29*** (0.129)	12.33*** (0.138)
Masters dummies	NO	NO	NO
Constrained	YES	YES	YES
Masters dummies			
Characteristics	NO	YES	YES
Father occupation	NO	NO	YES
Mother occupation	NO	NO	YES
Observations	1,316	1,310	1,310
R-squared			

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

The absence of a large sorting across masters is confirmed by two other facts. First, there is not much difference in selectivity across master degrees, with the first year results averages between masters being very close from each other, with the exception of the dual degree with a prestigious business school, as well as the degree in economics, which are highly selective. However, these two masters gather only a small number of students. As a result, even though there is a sorting, it is unlikely to be based on the comparative difficulty of the master degrees.

Second, the sorting across master degrees based on admission entrance is important only for three master degrees, in journalism, human resource and in communication. However, this masters are quite small, and gather only seven percent of the students. In the other hand, the masters in public affair and the PhD are less attractive for affirmative action. This points to the fact that these students are mostly attracted by the most vocational master degrees, and tend to avoid the master who require to pass difficult exams after graduation in order to obtain a secure job. All in all, while real,

the sorting based on the way of being admitted seems too small to bias the estimates, and should be eliminated by the master constrained fixed effects.

We present below the percentage of affirmative action students in each master degree, as well as the mean GPA of these degrees.

VARIABLES	(1) % Affirmative action students mean	(2) mean first year GPA mean
Public affair	0.058	13.28
European affair	0.073	13.33
International affair	0.089	13.13
Private Law	0.090	12.90
Public Law	0.071	13.32
Economics	0	13.84
PhD Program	0.038	13.05
Finance	0.154	12.98
Environment	0.028	13.11
Communication	0.165	12.47
Urban studies	0.135	13.21
Human resource	0.409	12.61
Marketing	0.256	12.71
Journalism	0.211	13.29
Dual degree BSchool	0	14.65
Total	0.107	13.09
Observations	1,316	1,316

To assess convergence, we also look at the ranking of students. We rank each student in the first year able to go until the fourth year compared to all members of his or her cohort (able to go to fourth year) based on average GPA. As for the fourth year, we first use a ranking based on the raw cohort GPA. Then we look only at the ranking within masters, and third, we remove master fixed effects (obtaining in regression controlling for first year results) to GPA before computing the ranking. In general these three rankings do not differ very much.

---

VARIABLES	(1) Rank first year	(2) Rank fourth year	(3) Rank fourth year (by master)	(4) Rank fourth year (adjusted)
Admission : Highest Honors	0.601	0.553	0.552	0.555
Admission : Science track	0.473	0.591	0.556	0.603
Admission : Math track	0.298	0.511	0.575	0.502
Admission : humanities track	0.449	0.598	0.523	0.510
Admission : Affirmative action	0.248	0.346	0.328	0.315
Admission Exam	0.473	0.476	0.486	0.483
Observations	1,317	1,317	1,317	1,317

---

We see again that there is a significant but incomplete convergence. In first years, the average rank of affirmative action students is about 0.24. In fourth year, it goes up to between 0.31 to 0.34 according to the different specification, which means that affirmative action students close about one third of the gap in term of rank with the other students.

### 3.4 Is the selection process rational ?

#### 3.4.1 The determinants of failure and convergence

There is some heterogeneity among the affirmative action students. As a result, it is interesting to see if some of their characteristics can predict failure, or in the opposite, convergence.

We first run logit regression to predict the probability of repeating the first year, pass with conditions (namely if some exams are retaken), and fail to graduate from the bachelor. Due to the small number of observations, we do not add control variables such as the characteristics of the peer groups. However, given that students are assigned at random, this should not matter.

VARIABLES	(1) Fail 1st year	(2) conditional passing	(3) Fail "premier cycle"
Cohort 2009-10	1.060** (0.412)	-0.154 (0.386)	1.158** (0.505)
HS humanities track	0.602 (0.478)	-0.346 (0.612)	1.158** (0.541)
HS Scientific track	-0.948* (0.548)	-0.177 (0.467)	-0.489 (0.648)
Other track in HS	2.121* (1.269)		2.594* (1.371)
Overseas HS degree	1.006 (0.651)	-0.721 (0.688)	1.889** (0.753)
HS degree obtained in Paris region	0.158 (0.520)	-0.684 (0.471)	0.247 (0.657)
Highest honor at HS graduation exam	-1.619** (0.659)	-2.739*** (1.050)	
Female	-0.258 (0.383)	-0.829** (0.387)	-0.169 (0.455)
Foreigner	0.708 (0.712)	-0.0697 (0.858)	0.806 (0.782)
Dual nationality	0.168 (0.563)	-1.248 (0.793)	-0.765 (0.826)
Constant	-1.825*** (0.607)	0.159 (0.523)	-2.684*** (0.786)
Observations	211	207	167

Standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Only a few numbers of coefficients are significant. Having passed the end of high school exam is consistently associated with better outcome. In fact, no student with such an achievement fail to graduate from bachelor, which compel us to remove them from the sample. Students from the most recent cohort performs less well, probably because they are less selected. Students having studied in the literary track have less chance to graduate, as well as the students from overseas territory, the effect being quite large. This is consistent with the hypothesis that only the best prepared students and those having the best credentials can succeed.

We then look at the determinant of convergence. Again, we have a very small sample of 142 students. The characteristics have a very small predictive power, with a R-square below 0.07, and a negative adjusted R-square. The only significant coefficient is a negative coefficient for student having graduated from high school with honor. This

is probably due to the fact that these students fare quite well in the first year of study, and that as a result, they are less likely to improve their rank.

VARIABLES Specification	(1) rank improvement (raw)	(2) rank improvement (master fixed effects)	(3) rank improvement (adjusted)
HS humanities track	0.0327 (0.0654)	0.0464 (0.0633)	0.0758 (0.0611)
HS Scientific track	-0.0291 (0.0479)	-0.0219 (0.0464)	-0.0155 (0.0448)
Other track in HS	-0.0819 (0.250)	-0.0201 (0.242)	-0.0675 (0.233)
Overseas HS degree	-0.106 (0.0891)	-0.0633 (0.0862)	-0.0911 (0.0833)
HS degree obtained in Paris region	-0.0403 (0.0527)	-0.0406 (0.0510)	-0.0322 (0.0493)
Highest honor at HS graduation exam	-0.100** (0.0498)	-0.0581 (0.0482)	-0.0886* (0.0466)
Female	0.0629 (0.0462)	0.0291 (0.0447)	0.0170 (0.0432)
Foreigner	-0.0304 (0.113)	-0.000346 (0.110)	-0.0415 (0.106)
Dual nationality	-0.108 (0.0653)	-0.101 (0.0632)	-0.100 (0.0610)
Constant	0.140** (0.0554)	0.122** (0.0536)	0.118** (0.0518)
Observations	142	142	142
R-squared	0.062	0.042	0.062

Standard errors in parentheses  
 \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

To conclude, we see that the university is not ignoring any factor which could predict a stronger convergence. However, it does not take into account some information about the likelihood of graduating from the bachelor.

### 3.4.2 The impact of affirmative action students on their peers

I then try to see if affirmative action students have an impact on the academic results of their peers. Indeed, it is often argued that one of the benefit of affirmative action is to improve the diversity of the students, and to create a better learning environment.

Conversely, since affirmative action students are in average weaker than their peers, they could exert a negative influence on them. In each specification, I control for all the characteristics available for each student. I then add the average of the characteristics of the other students of the group.

In this specification, I do not try to separately identify the exogenous and endogenous peer effects. Given than we have a random assignment of students to group, I can rule out bias coming from an endogenous sorting into groups, as well as correlated effects.

VARIABLES	(1) Results 1st year	(2) Results 1st year	(3) Results 1st year	(4) Results 1st year
Mean Admission : Highest Honors	0.322 (0.356)	0.746 (0.473)	0.733 (0.544)	0.727 (0.616)
Mean Admission : Affirmative action	0.424 (0.498)	0.565 (0.685)	0.292 (0.871)	-0.140 (0.986)
Constant	11.85*** (0.220)	13.45*** (1.208)	13.79*** (1.294)	14.38*** (1.378)
Peer characteristics	NO	YES	YES	YES
Peer Father occupation	NO	NO	YES	YES
Peer Mother occupation	NO	NO	NO	YES
Characteristics	YES	YES	YES	YES
Father occupation	YES	YES	YES	YES
Mother occupation	YES	YES	YES	YES
Observations	1,595	1,595	1,595	1,595
R-squared	0.328	0.333	0.334	0.339

Standard errors in parentheses  
 \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

We see that there is consistently no effect of being with a larger number of peers benefiting from affirmative action. In general, no one of the coefficients are significant. This is consistent with other studies who found no peer effect in higher education, and thus unsurprising.

### 3.5 Conclusion

This study analyzed an affirmative action program in France, based mostly on geographical and social origin rather than race or ethnic origin. It showed that, contrary

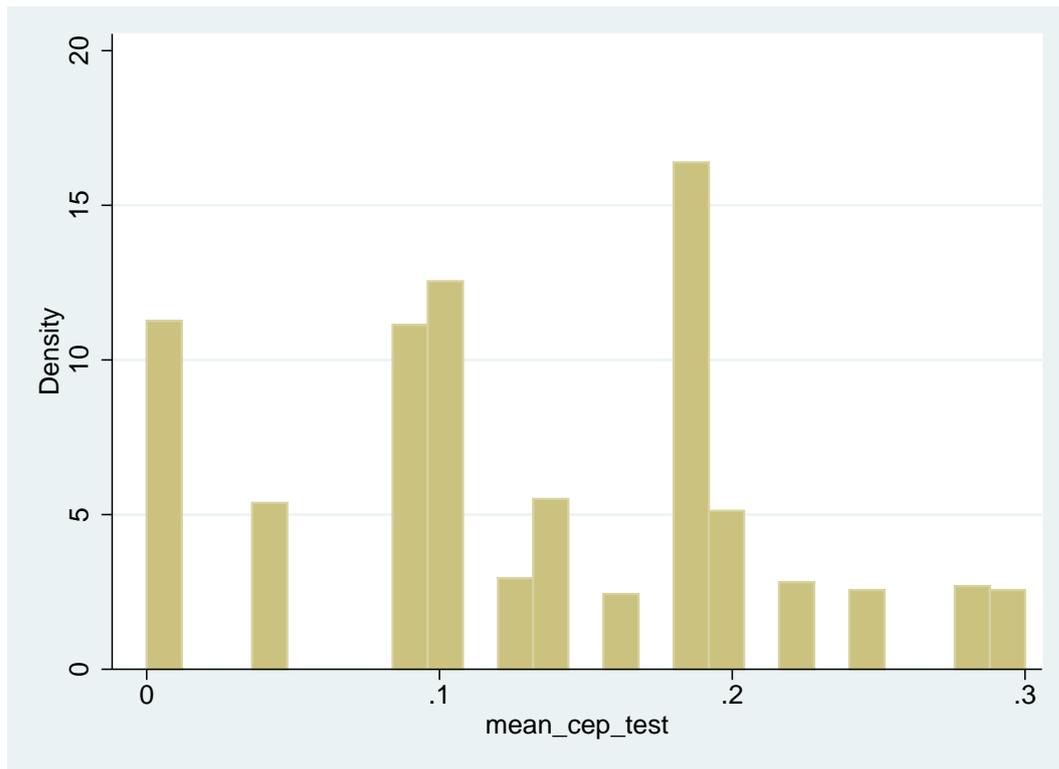
to the experience of many US programs, students benefiting from the program were able to partially catch up with their peers, conditionally on being able to go until their fourth years of study. However, they are also much more likely to fail. It also shows that there is only little sorting across majors, and that affirmative action recipients are more likely to choose more vocational majors. There is no good predictor of the ability to catch up. However, there are some good predictors of failure, that is high school track and being able to graduate from high school with honor.

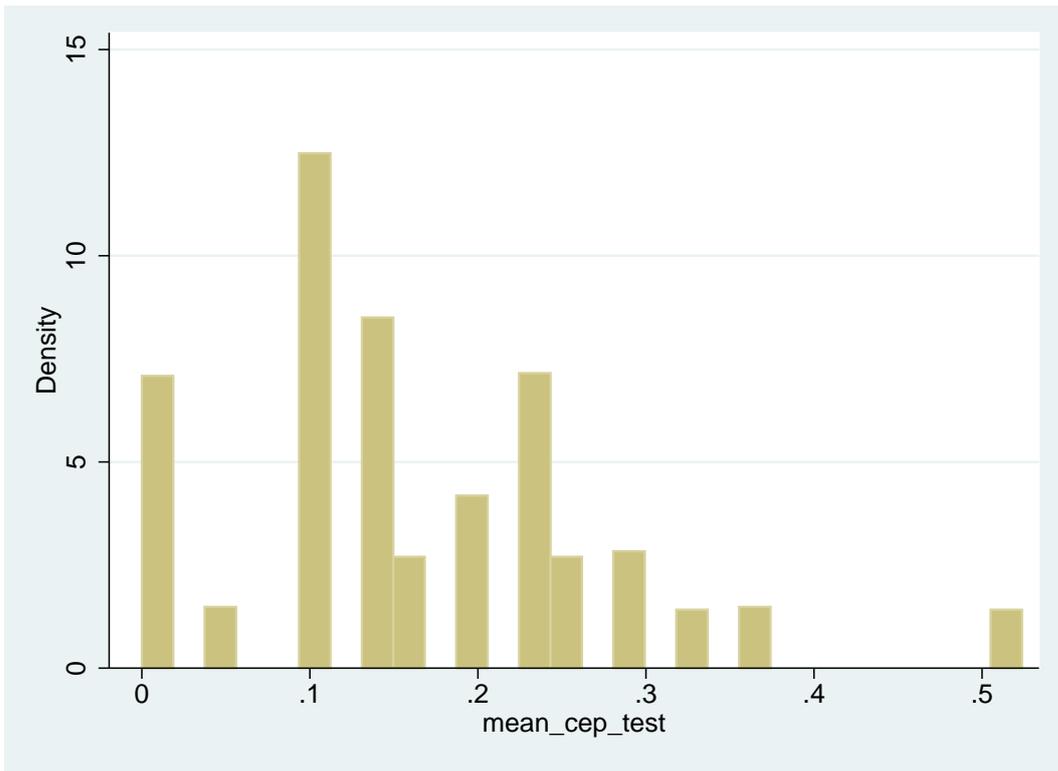
As a result, this study provides some valuable knowledge for policy purpose. First, affirmative action program based on social origin are more likely to exhibit catch up from their recipients than traditional programs based on raced and ethnic belonging. Second, the university is ignoring some predictors of failure while recruiting its students. It is not only true for affirmative action students, but also for the students entering through regular admission procedures. As a result, one can think that the university could improve the effect of the affirmative action program by better screening its applicants. In general, affirmative action program could benefit from selecting disadvantaged students who have more ability to improve academically.

## 3.6 Appendix

### 3.6.1 Appendix I : the allocation of affirmative action students across groups

We show below the distribution of the percentage of affirmative action students in groups, for the cohort 2007-2008 and 2009-2010.







## Troisième partie

### Les effets de pairs



## Chapitre 4

# L'économétrie des réseaux et des effets de pair

## 4.1 Introduction : l'importance des effets de pair pour un marché scolaire

Une question importante dans la discussion des conséquences de la création d'un marché de l'éducation est celle de l'ampleur des effets de pair. Les effets de pair correspondent à l'impact sur un individu du comportement et des caractéristiques des personnes avec qui il interagit. Il peut s'agir, par exemple, des camarades de classes, des membres d'un même club de sport, ou encore des habitants d'un même quartier. L'étude des effets de pair s'est beaucoup développée depuis les années 1990, et il semble que les pairs ont un effet sur les résultats scolaires, les comportements à risque (fumer, boire), et de nombreux autres comportements sociaux.

L'existence d'effets de pair est une question importante car la création d'un marché de l'éducation est susceptible de changer la répartition des élèves au sein du système éducatif. En effet, les élèves ne sont pas répartis aux hasards. Il existe une ségrégation relativement forte entre bons et mauvais élèves, élèves riches et élèves pauvres.

Cela est dû au fait que dans la plupart des systèmes scolaires existants, la fréquentation d'une école donnée est déterminée par la résidence dans un quartier voisin, en fonction d'une carte scolaire fixée administrativement. Le choix de l'école se fait donc de manière indirecte à travers le choix d'un quartier. Or, en général, les ménages les plus riches tendent à se regrouper géographiquement, donnant lieu à un phénomène de stratification urbaine. Cela est lié au fait que les ménages les plus riches ont la possibilité financière de résider dans les quartiers ayant le plus de caractéristiques attractives, et qu'ils valorisent également le fait de résider dans des quartiers où les ménages ont un niveau social égal ou supérieur au leur (voir par exemple l'ouvrage d'Eric Maurin, "Le ghetto français" [128], pour le cas de la France). Par conséquent, il existe également une forme de stratification scolaire, avec des établissements hiérarchisés en fonction de leur public scolaire (voir par exemple Epple et Romano, 1998, et Nechyba, 2000, pour des exemples de modèles de stratification résidentielle).

Le fait que la qualité des écoles dépendent des pairs, et que les meilleurs élèves soient également en moyenne plus riches renforce ce phénomène de stratification. On peut ainsi observer que la présence d'une bonne école augmente la valeur immobilière de sa zone de recrutement. Sandra Black (2000) [26] montre ainsi que le fait d'être situé de part et d'autre d'une ligne de démarcation entre deux districts scolaires a un effet non négligeable sur les prix. Fack et Grenet (2010) [61] documente un phénomène similaire à Paris.

La création d'un marché scolaire dérégulé est susceptible d'amplifier le phénomène de stratification scolaire dans la mesure où il peut permettre à des élèves doués mais pauvres d'accéder à de meilleures écoles. Si les écoles ont la possibilité de sélectionner les élèves selon leur niveau, voire de les discriminer au niveau des frais d'inscription en fonction de leur richesse et de leurs capacités dans un système de voucher (chèque édu-

cation), l'augmentation du choix scolaire accroît les écarts de niveau entre école ainsi que la ségrégation entre bons et mauvais élèves. Epple et Romano (1998) [60] ainsi que Nechyba (2000) [132] évalue les conséquences d'une libéralisation du marché scolaire à l'aide de modèle d'équilibre résidentielle, et trouve une augmentation importante de la stratification entre élèves à la suite de la disparition du zonage résidentielle.

Ce phénomène soulève donc deux difficultés. Tout d'abord, il crée un problème d'équité dans la mesure où les élèves les plus faibles souffrent de se retrouver avec des pair plus faibles, et inversement. Les écarts entre élèves tendent donc à augmenter, ce qui est problématique dès lors que l'on valorise un système scolaire égalitaire. Par ailleurs, les effets de pairs ne sont pas forcément linéaires. Il est possible que les élèves les plus faibles gagnent plus à la fréquentation d'élèves forts que les autres, et inversement. Dans le premier cas, une stratification accrue mène donc à une baisse des scores, alors qu'elle produit une augmentation des scores dans le second cas. Dès lors, estimer les effets de pair de manière précise, ainsi que leur possible non linéarité est important d'un point de vue de politique publique.

Cependant, mesurer les effets de pairs est difficile. On s'attachera donc tout d'abord à présenter les principales difficultés économétriques liées à l'estimation des effets de pair, ainsi que les méthodes proposées pour les résoudre. On présentera ensuite les principaux résultats empiriques obtenus. On discutera finalement des récents travaux sur les effets de réseaux, qui approfondissent la notion d'effets de pairs.

## 4.2 Effets de pairs et problèmes d'identification

A la suite de Manski (1993) [126], on peut distinguer différents types d'effet de pairs. Tout d'abord, les individus peuvent être affectés par les caractéristiques fixes ou antérieures de leurs pairs (par exemple, leurs résultats académiques passés, leur origine sociale). Il s'agit d'effets de pairs exogène. Par ailleurs, ils peuvent aussi être affectés par le comportement contemporain de leur pair, tel que leurs résultats académiques présents. Il s'agit d'effet endogène. En effet, le comportement des pairs est également influencé par l'individu lui même. Il existe donc une rétroaction. Enfin, L'ensemble d'un groupe de pair peut être influencé par des facteurs communs (par exemple, un professeur identique), qui peut être aisément confondu avec les effets de pairs. Il s'agit d'un effet corrélé.

Formellement, identifier les effets de pair revient donc à estimer l'équation suivante :

$$y_{ig} = X_i\beta + \frac{\lambda}{n_g} \sum_{j \in g} y_j + \frac{\delta}{n_g} \sum_{j \in g} X_j + \alpha_g + \epsilon_i \quad (4.1)$$

où  $y_{ig}$  correspond au résultat de l'individu  $i$  au sein du groupe  $g$  (par exemple ses résultats académiques, ou bien s'il fume, etc),  $X_i$  correspond à ses caractéristiques exogène

observées par l'économètre, et  $\alpha_g$  correspond à un effet fixe groupe qui comprend à la fois les effets corrélés au niveau du groupe et la corrélation inobservée entre les résultats des individus d'un même groupe lié par exemple à une formation endogène des groupes, et  $n_g$  est la taille du groupe. L'effet endogène est représenté par  $\lambda$ , et l'effet exogène par  $\delta$ .

Par conséquent, il existe un problème de réflexion lors de l'estimation des effets de pair. L'effet des caractéristiques fixes (les effets de pair exogènes) est amplifié par l'effet de rétroaction induit par les effets de pair endogène. Manski montre ainsi qu'il est impossible d'identifier séparément les effets de pair endogène et exogène en régressant les résultats contemporains sur une mesure des caractéristiques des pairs. Le problème de rétroaction devient évident dès lors que l'on transforme l'équation précédente en prenant la moyenne des groupes. On obtient alors la forme réduite suivante :

$$\bar{y}_g = \bar{X}_g \frac{\beta + \delta}{1 - \lambda} + \frac{\alpha_g}{1 - \lambda} + \frac{\bar{\epsilon}_i}{1 - \lambda} \quad (4.2)$$

On voit immédiatement qu'il n'est pas possible d'estimer séparément les coefficients du modèle. Dès lors que  $\lambda$  est positif et différent de zéro, les autres coefficients voient leur effet augmenter. On voit également que l'effet fixe groupe est colinéaire avec la moyenne des caractéristiques du groupe, et qu'il ne peut donc être estimé séparément de l'effet de pair endogène.

Cependant, l'existence d'effets corrélés crée un risque de biais dans la mesure où les groupes de pairs les plus favorables peuvent avoir plus de chance de recevoir des chocs négatifs ou positifs (par exemple, les enseignants les plus efficaces au sein d'une école pourraient être affectés systématiquement aux meilleures ou au moins bonnes classes). Une autre difficulté est liée à la possibilité d'un tri entre élèves et groupes de pairs fondé sur les inobservable. Ainsi, on peut imaginer que les élèves assignés aux meilleures classes partagent également des facteurs favorables qui ne sont pas observés. Formellement,  $\alpha_g$  est susceptible d'être corrélé avec  $\bar{X}_g$ , que cela soit à cause des effets corrélés ou de la formation endogène des groupes.

Enfin, la définition même des groupes de pair n'est pas évidente. En effet, il est rarement possible d'observer directement les limites des différents groupes en présence, qui sont de toute façon rarement hermétique. Ainsi, dans le cas du système scolaire, doit-on considérer que les pairs incluent les élèves d'une même école, ou bien d'une même cohorte, ou bien seulement d'une même classe ? Faut-il même considérer que les pairs correspondent uniquement au réseau d'amis ? En pratique, la définition a toujours une dimension arbitraire, qui a doit être gardé à l'esprit lors de l'interprétation des résultats. En effet, les effets de pairs peuvent être faible ou inexistant quand ils sont mesurés dans un groupe large, mais être beaucoup plus important au sein de sous-ensembles de ce groupe.

Les différents types d'effets de pair, ainsi que les groupes dans lesquels ils sont pré-

sents, ont un impact important en termes de politique publique. Ainsi, la présence d'effet de pair endogène implique l'existence d'un multiplicateur social. Un événement, telle qu'une nouvelle politique publique, qui affecte un individu dans un groupe a ainsi un impact sur l'ensemble du groupe à travers les effets de pair. A l'inverse, les effets de pair exogènes n'induisent pas ce mécanisme.

### 4.3 Les méthodes classiques : expériences naturelles, variables instrumentales et effets fixes

Différentes stratégies ont été proposées pour résoudre les problèmes d'identification exposés plus haut. Une stratégie simple consiste à estimer uniquement des formes réduites incluant à la fois des effets de pair endogène et exogène, sans chercher à résoudre le problème de la réflexion. Il ne reste plus alors qu'à résoudre le problème des effets corrélés, et de la constitution non aléatoire des groupes. Une première stratégie consiste à faire des hypothèses sur la manière dont les groupes sont constitués. Ainsi, Ammermueller et Pischke (2009) [7] font l'hypothèse qu'au sein d'une école donnée, les élèves sont alloués de manière aléatoire entre les classes, et que le problème de tri entre élèves existe uniquement au niveau de l'école. Par conséquent, il contrôle pour ce tri à l'aide d'effets fixes écoles, et estime ensuite directement l'effet de la moyenne des caractéristiques des élèves de chaque classe sur les résultats. Pischke et al. trouvent un effet assez important du niveau socio-économique des pairs sur la réussite.

Cette méthode repose donc sur l'hypothèse que le tri n'existe pas dans la formation de certains groupes. Si cette hypothèse peut être arbitraire, elle peut néanmoins être défendable dans le cas d'expériences naturelles bien identifiées. Beaucoup d'études reposent donc sur un événement assurant une allocation aléatoire des individus dans les groupes. On peut par exemple utiliser un contexte institutionnel où l'on cherche explicitement à mélanger les étudiants ou bien un événement entraînant un afflux de nouveaux individus dans les groupes (par exemple une catastrophe naturelle entraînant un afflux d'enfants réfugiés dans des écoles).

Il s'agit avant tout d'étude réalisée dans un contexte universitaire où les étudiants sont alloués au hasard dans des groupes d'études, des dortoirs, ou des chambres. Ainsi, Sacerdote (2001) [152] et Zimmerman (2003) [159] ont examiné les effets de pair entre étudiants partageant une même chambre ou un même immeuble à Dartmouth College et Williams College, et trouve des effets de pair modéré sur les résultats universitaires, et nuls sur les décisions d'orientation (choix d'un sujet académique). Stinebrickner et al. (2006) [157] trouve également des effets de pair faible à Berea College, et avance que l'assignement d'un compagnon de chambre ne mène pas forcément à des interactions sociales profondes afin d'expliquer ce fait. De manière similaire, Lyle (2007) [122] ne

trouve pas d'effet de pair important à Westpoint, la plupart des similitudes en termes de réussite des étudiants d'un même groupe s'expliquant par des facteurs affectant l'ensemble du groupe (ie. des effets corrélés). Foster (2006) [68] trouve des résultats identiques à l'université du Maryland. Cependant, Carrell et al. [40] trouvent des effets pair relativement large dans leur étude de l'US Naval Academy. Cela semble du au fait que dans ce cas, les interactions au sein du groupe sont extrêmement fortes, les étudiants d'un même groupe passant l'essentiel de leur temps d'étude ainsi que de leur temps récréatif ensemble. La définition des groupes et l'intensité des interactions sont donc importants.

Des études utilisant des assignements aléatoires ou quasi-aléatoires ont également été menées à plus grande échelle afin de mesurer les effets de pair au niveau du primaire et du secondaire. Ainsi, Imberman et al. (2013) [97] étudient l'impact de l'arrivée à Houston d'élèves défavorisés venant de la Nouvelle-Orléans à la suite de l'ouragan Kathrina. L'arrivée des réfugiés se fait de manière aléatoire, ce qui garantit l'identification. Ils trouvent des effets de pairs de taille modeste, avec la présence de non linéarité dans la mesure où les élèves les plus faibles ont un impact négatif beaucoup plus fort que les autres. Il confirme également la monotonie des effets de pair (avoir de meilleurs pairs augmentent systématiquement les résultats).

En Chine, Ding et Lehrer (2007) [55] ainsi que Carman et Zhang (2012) [39] utilise les particularités du système éducatif chinois pour identifier de manière convaincante les effets de pair. En effet, les élèves chinois sont assignés à un lycée et à une classe en fonction de leur rang, en suivant une règle administrative rigide, de manière à mélanger les étudiants sans créer de classes de niveau. Le problème de sélection endogène est donc inexistant. Par ailleurs, les auteurs de ces deux études disposent également d'information sur les enseignants, ainsi que sur les moyens attribués aux écoles, ce qui permet de contrôler en grande partie pour les effets corrélés. Dans ces deux études, les auteurs trouvent des effets de pairs significatifs. Kang (2007) [107] utilise une stratégie similaire en Corée du Sud, et trouve également des effets de pairs significatifs. Cependant, dans ces trois cas, les échantillons utilisés sont de taille relativement réduite.

Un autre moyen de limiter les problèmes d'endogénéité est de combiner l'utilisation d'effets fixes pour chaque élève et des chocs exogènes affectant l'allocation des élèves au sein du système scolaire.

Ainsi, Hanushek et al. (2003) [81] étudient les effets de pair dans les écoles élémentaires au Texas, mesuré au niveau de l'école plutôt que de la classe. Afin de limiter le problème de tri entre groupe et celui des effets corrélés, ils utilisent des effets fixes au niveau des niveaux scolaires, des écoles et des cohortes, ainsi que des effets fixes élèves. Cela est possible dans la mesure où ils disposent de données longitudinales. Les effets de pairs sont donc identifiés par la variation d'une année sur l'autre pour un même élève de son groupe de pair. Afin d'éviter le problème de la réflexion, les auteurs utilisent comme mesure des pair les scores obtenus deux années auparavant. Cependant, cela

ne pallie que partiellement ce problème dans la mesure où les groupes de pair ne varie que partiellement au sein d'une même cohorte dans la même école, et que le problème de réflexion existe déjà auparavant. Une autre difficulté est que les groupes de pair auxquels est confronté un élève dans une cohorte donnée varie relativement peu dans la mesure où une part importante des élèves reste dans la même école durant toute leur scolarité. Par conséquent, l'identification vient avant tout du changement d'élève entre école, et du changement de statut d'un élève au cours de sa scolarité. Cependant, ces changements peuvent en eux-mêmes être perturbateurs pour les élèves. De plus, ils peuvent être corrélés avec des facteurs liés à l'école tels que l'amélioration de l'enseignement. Enfin, les élèves transférés peuvent avoir des particularités inobservables biaisant les estimations. Finalement, Hanushek et al. trouvent un impact relativement fort des pairs sur la réussite, compris entre 0.15 et 0.20 % d'un écart type.

D'autres études utilisent des techniques relativement similaires se fondant sur des effets fixes pour contrôler pour l'hétérogénéité inobservée entre élèves et entre groupes. Lavy et al. (2012a) [113] exploitent le fait que les élèves anglais au collège sont testés dans trois disciplines différentes pour utiliser des effets fixes élèves calculé à travers les disciplines. Ils regardent donc l'impact sur les résultats dans une discipline d'avoir des élèves relativement fort dans cette discipline par rapport aux autres. En ajoutant des effets fixes par écoles et cohortes, ils peuvent ainsi contrôler de manière efficace pour le problème de tri entre élève. De plus, ils utilisent les scores à l'entrée au collège pour mesurer le niveau des élèves, ce qui permet d'éviter le problème de réflexion. Dans la mesure où il y a une forte redistribution des élèves entre le primaire et le secondaire, les scores à l'entrée au collège peuvent être considéré comme exogène. De cette manière, la seule source de biais restante est l'existence d'un tri entre élève se fondant sur leur différence de capacité entre chaque sujet, ce qui est assez improbable. Lavy et al. trouvent des effets de pair presque inexistant, sauf pour les élèves les plus faibles, appartenant au cinq derniers centiles de la distribution des scores, qui ont un large effet négatif sur les autres.

Les effets de pair semblant avoir un impact avant tout à travers l'exposition aux élèves les plus faibles, qui sont probablement aussi ceux ayant le plus de problèmes de comportement, certains chercheurs ont essayé d'identifier les élèves les plus perturbateurs et de mesurer leur impact sur les autres élèves. Ainsi, Lavy et al. (2012b) [111] trouvent un impact négatif des élèves ayant déjà redoublé sur les autres. Similairement, Carell et Hoekstra (2010) [41] trouvent un impact négatif de la présence d'élève victime de violence domestique, qui ont le plus de chance de se montrer perturbateur en classe. Figlio (2007) [65] trouvent des effets similaire de la présence d'élèves recevant un prénom féminin, eux aussi plus à risque de se montrer perturbateur.

Ainsi, au sein du système éducatif, les effets de pair semble se conformer au modèle développé par Lazear [114], dans lequel les élèves ont une influence dans la mesure où ils perturbent le temps d'apprentissage. Par conséquent, plutôt que les résultats des pairs,

c'est plutôt leur comportement qui importe, et plus particulièrement, celui de quelques perturbateurs qui ont la capacité d'interrompre le processus d'apprentissage.

Certains travaux se sont intéressées à d'autres dimensions des effets de pairs, essayant en particulier de mesurer l'impact des pairs sur les comportements délinquants. Ainsi, Billings et al. (2013) [24] mesurent l'impact du taux de minorité ethnique sur les résultats scolaires et les comportements déviants en utilisant une variation quasi-aléatoire de la composition des écoles dans un comté américain. Ils exploitent la fin d'un programme d'allocation égalitaire des étudiants fondé sur leur origine ethnique due à une décision de justice. Ce système est remplacé par un zonage fondé sur le lieu de résidence, qui augmente très fortement la ségrégation entre groupes. Afin d'éviter les problèmes liés à des différences inobservables entre les étudiants de chaque école, Deming et al. comparent des élèves habitant dans des quartiers identiques, mais situés de part et d'autres de la limite entre deux zones de recrutement. De cette manière, on peut comparer des étudiants quasi similaires mais faisant face à des groupes de pair très différents. Deming et al. trouvent des résultats nuls ou très faiblement négatifs en ce qui concerne les résultats aux test académiques. Par contre, ils trouvent des résultats relativement large en ce qui concerne le taux de diplômation du lycée et la probabilité de s'inscrire dans une institution universitaire pour les étudiants du groupe dominant, et la probabilité d'être arrêté par la police pour les étudiants masculin venant du groupe minoritaire. Ainsi, une hausse de 10% du taux de minorité ethnique dans un établissement scolaire donne lieu à une baisse d'environ deux point de pourcentage de la probabilité d'être diplômé du lycée et de poursuivre des études universitaires pour les étudiants venant du groupe majoritaire, et à une hausse d'environ 1,5 point de la probabilité d'être arrêté par la police pour hommes issu d'une minorité ethnique. Etant donné que le taux de minorité ethnique dans un établissement varie entre 20 à 80%, ces effets peuvent être considérés comme larges.

Similairement, Bayer et al. [21] étudient l'impact des pairs sur la récidive d'adolescent primo-délinquant envoyés en prison en Floride, en utilisant l'évolution aléatoire des camarades de cellules liés à la sortie et à l'entrée en prison des prisonnier.s Ils trouvent un effet relativement large sur la récidive dès lors que les partenaires de cellules ont commis des crimes similaires à ceux pour lesquels l'individu observé a été condamné.

Enfin, au delà de la qualité des groupes de pairs, certains auteurs se sont intéressés à la quantité de lien amicaux dans l'environnement de l'élève, et en particulier, au nombre d'amis d'un élève donné. Ainsi, Lavy et Sand (2012) [112] exploite une particularité du système éducatif de Tel-Aviv, qui demande aux élèves de faire la liste des amis avec qui ils souhaiteraient continuer à étudier au collège lors de la transition depuis l'école primaire. En utilisant les données ainsi collecté, Lavy et Sand peuvent mesurer le nombre d'ami qu'un individu possède dans sa classe. En comparant la richesse du

réseau social entre la classe du primaire et celle du collège, ils peuvent alors estimer l'impact de l'importance du réseau social de l'étudiant sur les résultats scolaires.

Lavy et Sand montrent que la présence d'amis réciproques (ie. qui se citent mutuellement comme amis) a un impact positif sur les résultats, ainsi que la présence d'amis "suiveur" (ie. qui citent l'individu comme amis, mais qui ne sont pas cités). Par contre, la présence d'amis non-réciproques (ie. l'inverse des "suiveur") a un effet négatif, mais très faible. Ces résultats font échos à d'autres travaux qui montrent que la présence d'une séparation entre école primaire et collège a un effet négatif sur les résultats, probablement à cause de la perturbation des réseaux amicaux qu'elle cause (Rockoff et Lockwood, 2010 [140]).

## 4.4 L'approche par les effets fixes

Les études mentionnées jusqu'à présent s'appuient principalement sur l'existence d'expériences naturelles assurant une formation exogène des groupes. Cependant, de telles circonstances sont rarement disponibles, et beaucoup de bases de données ne se prêtent donc pas à une telle approche. Arcidiacono et al. [15] proposent donc une approche plus générale qui nécessite seulement de disposer de données longitudinales.

Leur approche consiste à utiliser le caractère longitudinal des données pour estimer des effets fixes étudiants, ceux-ci étant observés à plusieurs reprises. Par la suite, ils utilisent la moyenne des effets fixes des membres d'un groupe comme mesure de la qualité des pairs dans ce groupe. Le caractère longitudinal des données permet également d'ajouter des effets fixes années, des effets fixes cohortes, et des effets fixes professeurs. Cette technique permet donc de contrôler de manière très flexible pour les différentes influences susceptibles d'affecter le résultat étudié. Par ailleurs, dans la mesure où les caractéristiques des pairs sont mesurées à partir d'effets fixes qui sont invariants, elles ne sont pas exposées au problème de la réflexion. De plus, les problèmes d'effets corrélés et de formation endogène des groupes sont éliminés par l'usage des effets fixes, qui sont mesurés sur plusieurs périodes, ainsi que par la possibilité de contrôler pour les influences de la cohorte et des professeurs. Il est également possible de tester la non linéarité des effets de pair en régressant différent quantile du groupe plutôt que la moyenne.

Formellement, Arcidiacono et al. estime l'équation suivante,

$$y_{igpct} = X_{it}\beta + \frac{\delta}{n_g} \sum_{j \in g} \alpha_j + \alpha_i + \xi_p + \mu_c + \epsilon_i \quad (4.3)$$

où  $\alpha_i$  correspond à l'effet fixe de l'étudiant  $i$  estimé sur plusieurs périodes  $t$ , tout en contrôlant pour l'effets de professeurs grâce à l'effet fixe professeur  $\xi_p$  et pour l'effet de la cohorte grâce à l'effet fixe cohorte  $\mu_c$ . L'influence des pairs est donnée par l'impact

de la moyenne des effets fixes des autres étudiants du groupes,  $\frac{1}{n_g} \sum_{j \in g} \alpha_j$ , à savoir  $\delta$ . Un inconvénient de cette méthode est qu'elle est computationnellement difficile à mettre en œuvre. En effet, elle nécessite de calculer simultanément des effets fixes sur plusieurs dimensions, ce qui implique que les techniques classiques de soustraction de la moyenne ou de différentiation ne peuvent être appliquées. Arcidiacono et al. proposent un algorithme pour estimer ce modèle à l'aide de la méthode des moindres carrés non-linéaires et de l'algorithme EM.

Burke et Sass (2013) [36] utilisent cette méthode sur une base de données administrative américaine regroupant un grand nombre d'enseignants et d'étudiants suivis sur plusieurs années. Ils trouvent en moyenne un effet faible des pairs dans le cadre du modèle linéaire en moyenne. Cependant, cela cache une hétérogénéité importante entre groupe. Le rôle des pairs tend à augmenter avec l'âge, et surtout, il affecte positivement bien plus les bons que les mauvais élèves, qui tendent à ne pas être affecté par l'exposition à des camarades meilleurs qu'eux, voire, à en souffrir s'ils sont exposés à des très bon élèves. On retrouve donc l'impact différencié des étudiants en fonction de leur niveau déjà mis à jour par Lavy et al.

## 4.5 Effets de pair et réseaux

Les principales études sur les effets de pair échouent donc la plupart du temps à trouver des effets de pair significatif, en particulier dans le système éducatif. Cependant, il est possible que cela soit du à une mauvaise définition du groupe de pair. En effet, la plupart des ces études font l'hypothèse que les effets de pair ont lieu au niveau du groupe pris dans son ensemble, par exemple au niveau de la classe. Par conséquent, elles supposent que chaque élève affecte l'ensemble du groupe par son comportement. Cela correspond bien à des modèles d'interaction au sein d'un groupe tel que celui développé par Lazear, où chaque élève perturbe le fonctionnement de toute la classe par son comportement. De fait, dans la plupart des cas où l'on trouve des effets de pair significatif, il s'agit de l'effet d'élèves faibles où susceptible d'avoir des problèmes comportementaux.

Cependant, on peut également envisager les effets de pair comme l'impact d'interaction proche et répété au sein d'un petit groupe d'ami, de taille inférieure à la classe. Dans ce cas, il s'agit d'un effet réseau plus que d'un effet groupe ou d'un effet classe. Au sein d'un groupe chaque étudiant possède son groupe de pair spécifique.

Le premier modèle économétrique cherchant à exploiter cette opportunité est celui développé par Lee (2007) [117]. Il utilise le fait que chaque étudiant au sein d'une même classe est exposé à un groupe de pairs légèrement différent de celui des autres étudiants, en raison du fait qu'il influence les autres étudiants, mais que lui-même ne fait pas partie de son propre groupe de pairs. Pour peu qu'il y ait suffisamment de variation dans la taille des différents groupes, il devient possible d'estimer un modèle

où sont explicitement présent effets de pair endogènes et exogènes. Il est aussi possible d'ajouter des effets fixes pour chaque classe, ce qui permet d'éliminer les biais induits par la formation endogène des groupes, ainsi les effets corrélés.

Formellement, la méthode de Lee revient à estimer l'équation suivante,

$$Y_g = X_g\beta + \lambda W_g Y_g + W_g X_g \delta + \alpha_g + \epsilon_g \quad (4.4)$$

où  $g$  dénote le groupe.  $Y_g$  correspond donc au vecteur des résultats des individus du groupe,  $X_g$  à la matrice de leurs caractéristique,  $\alpha_g$  à l'effet fixe groupe, tandis que  $W_g$  correspond à la matrice d'interaction entre les membres du groupe, similaire à la matrice utilisé pour décrire un graphe ou un réseau. Les éléments de cette matrice de dimension  $n \times n$  prennent une valeur égale à  $\frac{1}{n}$  quand les deux individus sont connectés, et zéro autrement. Dans le cas du modèle proposé par Lee, l'ensemble des éléments de la matrice d'interaction sont égaux à un, à l'exception des éléments de la diagonale (en raison de l'exclusion de l'individu dans le calcul de la composition des groupes de pairs).

L'intérêt de ce modèle est donc de proposer une estimation direct de l'effet endogène, et donc de l'effet multiplicateur associé à un choc affectant des élèves du groupe. Lee propose un estimateur par variable instrumentale, et un autre fondé sur le maximum de vraisemblance pour estimer ce modèle. Cependant, l'identification est faible, car reposant sur des variations très ténues au sein de chaque groupe. Par conséquent, l'estimateur converge lentement, est très peu efficace lorsque les groupes sont de taille importante. Les applications de cet estimateur sont donc plutôt décevantes (Boucher et al. 2013 [29]), la variance de l'estimateur étant trop large pour trouver des résultats intéressants.

Il est donc nécessaire de mieux connaître les réseaux au sein du groupe afin de disposer d'une variation plus importante pour chaque individu, ce qui permet d'obtenir une identification beaucoup plus forte. Cela offre également l'espoir de trouver des effets de pairs plus substantiels, dans la mesure où le groupe de référence est mieux défini. Formellement, cela revient donc à utiliser une matrice  $W_g$  qui est beaucoup plus "éparse", c'est à dire qui comprend un grand nombre de zéro et donc un relativement faible nombre d'individu connecté.

Bramoullé et al. (2009) [31] identifient les conditions d'identification d'un tel modèle, notamment les conditions que doit satisfaire la matrice d'interaction, et proposent un exemple d'estimation à l'aide des doubles moindres carrés (2SLS). Il est possible d'obtenir des résultats relativement précis même pour des échantillons de taille relativement modérés. Bramoullé et al. confirme par des simulations Monte-Carlo que l'estimateur est d'autant plus efficace que la matrice d'interaction comprend peu de liens. Ils rejoignent les résultats de Lee (2004) [116] qui explore les propriétés asymptotiques des modèles d'économétrie spatiale (qui inclue eux aussi des matrices d'interaction) et qui montre également que la convergence ainsi que l'efficacité des estimateurs de ces mo-

dèles dépendent du caractère éparé de la matrice d'interaction spatiale.

De tels modèles d'interaction ont été estimés sur des bases de données récentes. Ainsi, Lin (2007) [120] étudie l'impact des pairs sur les résultats scolaires à l'aide de données de l'enquête Addhealth, qui suit environ 50,000 adolescents américains sur plusieurs années, et fournit des données détaillées sur leur réseau d'ami. Il trouve un effet endogène relativement large, à environ 20% d'un écart type de note, ce qui est non-négligeable. Plusieurs caractéristiques exogènes des pairs ont également des effets sur la réussite scolaire.

L'économétrie des réseaux souffrent cependant de nombreux problèmes, exposé par Goldsmith-Pinkham et Imbens (2013) [79]. Ceux-ci font valoir que les réseaux ne sont pas forcément exogènes. Cela entraîne deux conséquences. Tout d'abord, les régressions de réseaux peuvent être biaisées dans la mesure où l'effet réseau qui est estimé peu en réalité est dû à une similitude inobservée entre deux individus connectés. Ce biais est analogue à celui lié à la formation endogène des groupes, mais est présent maintenant au niveau des pairs d'individu. Par ailleurs, si la formation du réseau est endogène, les prédictions en terme de politique publique change drastiquement. Par exemple, une intervention changeant une caractéristique d'un groupe d'individu peut en affecter non seulement en raison des effets de pair, mais aussi en raison de la modification de l'architecture du réseau qu'elle entraîne.

Par conséquent, l'existence de réseaux endogènes rend plus complexe de manipuler les effets de pairs, dans la mesure où manipuler la composition peut affecter les réseaux sous-jacents de manière imprévue. Ainsi, Carrell et al. (2013) [42] montre comment une intervention visant à modifier la composition des groupes d'étude dans une université américaine afin de promouvoir l'égalité entre étudiants a des effets opposés en raison du fait que les étudiants les plus forts tendent à se regrouper entre eux à l'intérieur d'un même groupe, excluant les étudiants les plus faibles de leur réseau.

Un autre problème évoqué par Goldsmith-Pinkham et Imbens est que les réseaux sont souvent mal mesurés, dans la mesure où beaucoup de liens ne sont pas déclarés, et que la force des liens n'est pas évaluée. On peut aussi ajouter que la littérature sur les réseaux se concentre uniquement sur les liens positifs, et pas sur les interactions négatives.

Goldsmith-Pinkham et Imbens ainsi que Lee and Hsieh (2013) [95] propose donc de manière indépendante un modèle économétrique capable de corriger pour l'endogénéité des liens amicaux. Ce modèle se rapproche d'une forme d'instrumentation des liens amicaux, une équation supplémentaire étant ajoutée pour modéliser la formation du réseau. Ainsi, Hsieh et Lee propose la spécification suivante,

$$\psi_{ij} = c_i\gamma_1 + c_j\gamma_2 + c_{ij}\gamma_3 + \sum_k \delta_k |z_{ik} - z_{jk}| + \epsilon_{ij} \quad (4.5)$$

ou  $\psi_{ij}$  est la valeur latente déterminant si  $i$  et  $j$  sont amis,  $c_i$  représente les caractéristiques observable de  $i$ ,  $c_j$  celle de  $j$ , et  $c_{ij}$  les caractéristiques propres à la dyade  $ij$ .  $|z_{ik} - z_{jk}|$  représentent les  $k$  variables inobservées par l'économètre, mais qui peuvent affecter la formation des liens amicaux. Elles sont introduites dans l'équation sous forme de la valeur absolue d'une différence entre les caractéristiques inobservées de deux individus pour refléter le fait que la formation des réseaux est fortement affectée par l'homophilie. Goldsmith-Pinkham et Imbens propose une équation relativement proche, mais moins générale afin de rendre l'estimation du modèle plus aisé.

Dès lors que  $|z_{ik} - z_{jk}|$  a à la fois un impact sur la formation du réseau, et que  $z_{ik}$  a un impact sur le résultat final  $y_i$ , on a un problème d'endogénéité : une régression de réseau naïve surestimerait l'impact des caractéristiques du réseaux en ne contrôlant pas correctement pour  $z_{ik}$ . Il est donc nécessaire de proposer une stratégie permettant de corriger pour ce problème d'endogénéité. Une manière intuitive de procéder est d'estimer de manière jointe l'équation finale et l'équation déterminant le réseau, si possible en utilisant des instruments dans l'équation déterminant la formation du réseau.

Rogowsky et Sinclair, deux politologues, proposent ainsi d'instrumenter l'existence d'un lien entre deux membres du congrès des Etats-Unis par la proximité de leurs bureaux respectifs au sein du capitol, qui est déterminée de manière aléatoire (2012) [143]. Contrairement à la plupart des études observationnelles sur les réseaux, ils ne trouvent aucun effet du réseau sur des comportements tels que le vote ou le soutien à des propositions de loi. Leur étude montre donc que les problèmes d'endogénéité des réseaux peuvent être important. Néanmoins, il est en général très difficile de trouver un bon instrument, ce qui rend difficile l'usage systématique de ce type de stratégie.

Hsieh and Lee propose une approche plus générale, fondée sur l'estimation jointe de l'équation déterminant les résultats et de celles déterminant les réseaux. Il discute l'identification d'un tel modèle, et montrent que l'usage d'instruments n'est pas nécessaire, et que le modèle peut être identifié uniquement grâce à des restrictions sur la distribution des variables inobservés. L'identification est cependant complexe, et se rapproche de l'idée des modèles de sélection tels que ceux développés par Heckman, qui reposent sur un choix spécifique de forme fonctionnelle.

Il n'est pas possible d'estimer le modèle précédent par maximum de vraisemblance, en raison de la complexité de la fonction de vraisemblance, qui inclut des éléments inobservés qui sont difficiles à éliminer par intégration. Par conséquent, aussi bien Hsieh et Lee que Goldsmith-Pinkham and Imbens utilisent des méthodes bayésiennes, basées sur la méthode MCMC. Tous deux appliquent leur modèle sur les données AddHealth, et trouvent des effets de pair significatifs, bien que faibles. Ils trouvent également des signes d'endogénéité, mais de magnitude relativement limitée.

## 4.6 Les modèles structurels de réseaux

Une autre manière de résoudre le problème d'endogénéité de la matrice d'interaction est d'adopter une approche structurelle et d'utiliser la théorie économique pour spécifier un modèle de formation de réseaux. Les économistes se sont beaucoup intéressés à la formation des réseaux depuis une vingtaine, en développant notamment des modèles stratégiques de formation de réseaux où les individus choisissent de former des liens amicaux en fonction des coûts et des bénéfices que peuvent générer ces liens (voir par exemple Jackson et Wolinsky, 1996 [99], et plus récemment Ballester et al., 2006 [20], et Bramoullé et al., 2014 [32]). Plusieurs études ont essayé d'évaluer empiriquement les prédictions de ces modèles, notamment les travaux de Fafchamps et Gubert (2007) [62], Calvo-Armengol et al (2009) [37].

Dans la mesure où cette littérature n'estime pas directement l'impact des amis sur les comportements et résultats individuels, elle n'est pas directement connectée à la littérature sur les effets de pairs. Par exemple, Calvo-Armengol et al. montre qu'un individu relativement central dans un réseau a plus d'intérêt à fournir d'effort en raison de la rétroaction positive qu'induit son impact sur ses relations et sur le réseau en général, et valident ces prédictions sur les données AddHealth. Fafchamps et Gubert étudient les déterminants des liens d'entraide dans une région rurale des Philippines et montrent qu'ils sont principalement dus à des facteurs contingents, tels que la proximité géographique ou familiale, plutôt qu'à des facteurs stratégiques tels que la recherche d'une diversification des liens assurantielles.

Cependant, des chercheurs ont essayé d'exploiter les enseignements de la littérature théorique sur les réseaux pour construire et estimer des modèles structurels de formation de réseaux et d'interaction sociale. Ainsi, Hsieh et Lee (2013) [96] s'inspire de la littérature sur la formation stratégique des réseaux pour proposer une équation plus riche modélisant la formation de liens sociaux. L'ensemble de l'utilité retirée par l'individu  $i$  de la participation au réseau  $W$  est donné par l'équation suivante :

$$\xi_i = \sum_j w_{ij}(c_i\gamma_1 + c_j\gamma_2 + c_{ij}\gamma_3) + \bar{\omega}(w_{i.}, W_{-i.})\eta + \sum_d \frac{\zeta_d}{2} y_{id}^2(W) \quad (4.6)$$

la première partie de l'équation reprend l'essentiel de l'équation (5), moins la partie inobservée. La seconde partie  $\bar{\omega}(w_{i.}, W_{-i.})\eta$  modélise l'utilité retirée du fait de la structure du réseau.  $\bar{\omega}$  est une statistique de réseau (par exemple, la mesure de centralité de l'individu  $i$ ) qui dépend à la fois des liens de  $i$  et du reste de la structure du réseau,  $W_{-i.}$ . Enfin, la dernière partie  $\sum_d \frac{\zeta_d}{2} y_{id}^2(W)$  correspond à l'impact des effets d'incitation. La forme spécifique que prend cet effet est dérivé du modèle de Ballester et al., qui calcule l'équilibre d'un jeu où chaque individu reçoit de l'utilité en fonction du comportement des autres membres de son groupe, et choisit son effort en fonction de l'utilité directe qu'il en retire ainsi que de l'utilité dérivée de la modification du comportement de ses

pairs induit par son effort.

Hsieh et Lee introduisent ensuite cette équation dans un modèle de graphe aléatoire exponentielle. Ce modèle mesure la probabilité d'apparition d'un graphe spécifique à partir d'un groupe donné à l'aide d'un modèle polynomial :

$$P(W_g) = \frac{\exp(Q(W_g))}{\sum_{W \in \Omega_g} \exp(Q(W))} \quad (4.7)$$

Ce modèle dépend donc d'une statistique de réseau  $Q(W_g)$ . Afin de construire cette statistique de réseau, Hsieh and Lee font l'hypothèse que le réseau est formé à l'aide d'un jeu coopératif. Par conséquent, le réseau obtenu à l'équilibre est celui qui maximise la somme des utilités définies dans l'équation (6), et dénoté par  $TU(W - g)$ . En ajoutant une perturbation  $\xi_{W_g}$ , on peut ensuite relier cette statistique au modèle défini dans l'équation (7), en remplaçant  $Q(W)$  par  $TU(W - g)$ . En effet, l'équation (5) donne la probabilité que  $Q(W_g) + \epsilon_{W_g}$  soit plus élevé que les statistiques correspondant aux autres réseaux possibles.

On obtient donc un modèle qui en théorie est estimable par maximum de vraisemblance, de même que tout modèle polynomial. Etant donné la très grande dimension de la fonction de vraisemblance, Hsieh et Lee utilisent des méthodes bayésiennes pour estimer le modèle.

Le modèle de Hsieh et Lee montre donc comment on peut utiliser la théorie économique pour spécifier un modèle de formation de réseau et d'effet de pairs. Si la structure du modèle est jugée crédible, on peut dès lors considérer que le problème de la formation endogène des liens est résolu.

De même que dans les études corrigeant pour l'endogénéité des réseaux, Hsieh et Lee trouve des effets de pair plus faible. De plus, ils sont capable d'identifier des paramètres économiquement intéressant décrivant comment les réseaux sont formés. Ils montrent ainsi que des considérations stratégiques liées aux résultats scolaires entrent en compte dans le cas de la formation des réseaux, les étudiants sélectionnant en partie leurs amis en fonction de leurs bons résultats scolaires et de l'impact positifs que cela peut avoir sur leurs propres efforts et résultats. Par contre, de tels effets d'incitation sont absents pour le fait de fumer. Les étudiants fumeurs ne cherchent pas particulièrement à être amie avec des personnes qui fument également.

## 4.7 Estimer les effets de réseaux sans connaître la matrice d'interaction

Jusqu'à présent, nous avons fait l'hypothèse que la matrice d'interaction est connue, même si elle est susceptible d'être endogène et observée avec erreur. Cependant, il est fréquent de ne pas posséder autant d'observation, notamment parce que rassembler

des données sur les interactions sociales est couteux et requiert l’usage de sondages. Par conséquent, des recherches récentes ont essayé de proposer des méthodes capable d’estimer les effets de pair et la matrice d’interaction simultanément.

Manresa (2013) [125] développe une méthode fondée sur l’estimateur du LASSO et permettant d’estimer le réseau en même temps que les effets de pairs. Utiliser le LASSO permet d’estimer des modèles de grande dimension, quand il y a un large nombre de régresseurs comparé au nombre d’observation. Dans ce cas, utiliser des méthodes standards telles que la régression linéaire est souvent impossible. Par exemple, supposons que l’on veuille mesurer l’impact de chaque individu sur les autres dans un réseaux inobservés en incluant une variable pour chacun d’entre eux dans le modèle. Dans ce cas, pour un réseau comprenant  $N$  individus, il y a  $N$  régresseurs, en plus des autres variables explicatives. Dès lors, il n’est pas possible d’estimer le modèle par OLS dans la mesure où il y a plus de régresseurs que d’observations.

Le LASSO permet d’estimer de tels modèles en sélectionnant un nombre limité de régresseur et en fixant les autres régresseurs à zéro. Pour cela, il minimise la somme des carrés et une pénalité augmentant avec le nombre de régresseur. Par conséquent, le modèle sélectionne de manière optimale les régresseurs qui diminuent le plus la somme des carrés, jusqu’au moment où le régresseur marginal diminue moins la somme des carrés qu’il n’augmente la pénalité. Calculer l’estimateur nécessite néanmoins de résoudre un problème d’optimisation complexe, et donc couteux en calcul, ce qui le rend difficile à utiliser sur de larges bases de données. Manresa propose donc le modèle suivant,

$$y_{it} = \alpha_i + \beta_i x_{it} + \sum_{j \neq i} \gamma_{ij} x_{jt} + \theta' w_{it} + \delta_t + \epsilon_{it} \quad (4.8)$$

où  $y_{it}$  correspond au résultat de l’individu  $i$  à l’instant  $t$ ,  $\alpha_i$  est un effet fixe capturant l’effet des caractéristiques invariantes de l’individu  $i$ ,  $\beta_i x_{it}$  représente l’impact des caractéristiques de  $i$  qui varie dans le temps, et  $\sum_{j \neq i} \gamma_{ij} x_{jt}$  correspond à l’impact de l’individu  $j$  sur  $i$ , qui dépend des caractéristiques de  $j$  à l’instant  $t$ , ainsi que d’un paramètre ne variant pas dans le temps,  $\gamma_{ij}$ , qui capture la force du lien entre  $i$  et  $j$ . Le résultat est également déterminé par des facteurs affectant tous les individus,  $\theta' w_{it}$ , qui représente l’impact de facteur commun à tous les individus (et dont l’impact  $\theta_i$  ne varie pas entre individus), et  $\delta_t$ , qui correspond à des chocs affectant tous les individus à un moment dans le temps.

Manresa propose un version modifiée du LASSO, où la fonction objectif à minimiser correspond à la somme des carrés du modèle, et à une pénalité augmentant avec le nombre d’interaction  $\gamma_{ij}$  qui sont positifs. Par conséquent, l’estimateur sélectionne quelles sont les liens entre individus qui sont positifs, et fixe les autres à zéro. De cette manière, il permet à la fois de révéler la structure des interactions, et de mesurer la force de chacune de ses interactions.

Une critique possible est que le modèle de Manresa n’inclut pas l’existence d’effet de pair endogène. Par conséquent, pour peu que ceux-ci soient différent de zéro, il est

vulnérable au problème de la réflexion. Cependant, cette lacune est en passe d'être comblée par les travaux récents de Rose (2014) [145], qui étend le modèle de Manresa en y incluant des effets de pair endogène, et qui propose une méthode de réflexion fondée sur l'estimateurs à variable instrumentale pour des modèles à grande dimension développé par Gautier et Tsybakhov (2013) [75].

Les travaux de Manresa et de Rose constituent donc une avancée majeure, dans la mesure où ils permettent d'estimer explicitement le réseau d'interaction entre individus. Cependant, ils comportent des limitations qui les rendent difficile à appliquer à l'estimation des effets de pair. Tout d'abord, l'estimateur de Manresa n'est applicable que sur des données de panel, avec un panel relativement long. L'estimation est coûteuse en terme de calcul, et ne peut être employé que sur des données de taille limitée. Ainsi, Manresa applique son estimateur sur un panel d'environ 300 firmes observées pendant une vingtaine d'année, afin de mesurer l'existence de retombée en terme de productivité des autres entreprises des dépenses en R&D d'une firme donnée. Cependant, la plupart des bases de données disponible (par exemple dans le domaine de l'éducation) sont beaucoup plus large, et n'ont qu'une dimension temporelle relativement réduite, avec des groupes qui restent rarement fixes au cour du temps. Les travaux de Manresa ne reste donc applicable qu'à des contextes très particuliers. Par conséquent, des progrès restent à faire pour que les nouvelles méthodes d'estimation de la structure des réseaux puissent être appliquées aux effets de pairs.

## 4.8 Conclusion

L'estimation des effets de pairs et des effets de réseaux présentent de nombreuses difficultés, qui ont donné lieu au développement d'une riche littérature théorique aboutissant à la mise au point de nouveaux outils économétriques. Ces outils ont permis d'obtenir un certain nombre de résultats convaincants. En particulier, il semble que les effets de pairs de nature cognitive soit réel, mais que leur magnitude soit faible. Ils semblent néanmoins être plus fort pour les individus les plus faibles, qui ont un effet négatif important sur les autres. Les effets de pairs concernant les comportements sociaux sont eux beaucoup plus large.

La recherche permet donc d'apporter une réponse nuancée à la question de l'existence d'effets de pairs, et donc de mieux guider les politiques publiques, au delà des leçons tirées des croyances communes.



## Chapitre 5

Measuring endogenous and exogenous  
peer effects using class size  
variation

## 5.1 Introduction

Many parents invest considerable time and efforts in order to ensure that their children are raised and schooled among good peers, that is, children coming from wealthy and educated families, and without behavioral and learning problems. As a result, peer effects are often believed to be an important parameter as far as educational and social policies are concerned. Thus, the impact of segregation, and the legitimacy of the measures aiming at ending it, crucially depends on the existence and shape of peer effects. Similarly, the magnitude of peer effects is crucial for assessing school choice policy which aims at broadening the set of school that a child can attend, and are often criticized for increasing segregation. Last, peer effects can have important implication on the optimal repartition of students in a school system.

However, while parents seem to strongly believe in a large impact of peer effects, empirical evidence are mixed. Many studies fail at finding policy significant academic peer effects. Thus, studies by Hanushek et al. (2003) [81] and Burke and Sass (2013) [36] found only a small impact of the average of the group academic results on student achievement for students in high schools. Similarly, studies on higher education find only very small peer effects (Sacerdote, 2001 [152], Zimmerman, 2003 [159], Lyle, 2007 [122], Foster, 2006 [68]). However, many studies find larger nonlinear peer effects, showing that the worst performing students, or some students having specific behavioral problems can harm the other, while other students have little impact (Carrel and Hoekstra, 2010 [41], Burke and Sass [36], Lavy et al., 2012a [113] 2012b [111], Imbermann et al., [97]). Moreover, studies focusing on the impact of peer effects on outcomes such that smoking or consuming drug found large effects (Billing et al., 2014 [24]). As a result, the existence, shape and size of peer effects is still an open question.

Estimating peer effects is made difficult by several econometric problems. First, as highlighted by Mansky (1993) [126], peer effects can be divided into endogenous peer effects (the effect of the contemporaneous behavior of other students) and exogenous peer effects (the effects of fixed characteristics of other students), which are in practice difficult to disentangle. However, failing to do so can lead to biased estimates of the effects of some characteristics on other people, due to the so-called reflexion problem. Indeed, the presence of endogenous effects means that there is a retroaction : by affecting his or her peers, a student will eventually impact his or her own behavior. As a result, a naive estimation of exogenous peer effect is likely to overestimate their importance.

Moreover, peer effects are easily confounded with correlated effect, the impact of common factors affecting all the members of a group (for instance, a common teacher). What's more, students tend to sort into groups based on unobservable characteristics, often correlated with observables, leading to spuriously attribute to peer effects some similarities across students which are not well captured by existing controls. All these

factors tend to bias upward estimates of peer effects.

Several strategies have been employed so far to overcome these problems. Some studies assume them away by focusing on group where sorting and correlated effect are unlikely to occur, and estimate reduced form coefficients including both endogenous and exogenous peer effects (Ammermueller and Pischke 2009 [7]).

Several studies have exploited random allocation of individuals to groups in order to solve both the sorting problem and the correlated effect problem (Sacerdote, Imberman et al, Zimmerman). Although natural experiments give good guarantees of identification, they can most of the time only estimate reduced form coefficients, except when they can instrument current achievement by lagged measures.

When natural experiments were not available, researchers have mostly used fixed effects approach, either trying to control for ability by using student fixed effects across subjects (Lavy et al. 2012a [113]), or across time (Arcidiacono et al. 2012 [15], Burke and Sass [36]). The approach by Arcidiacono et al. has the merit to solve the reflexion problem and the endogenous sorting problem. However, it estimates only the impact of a general measure of peer ability (their fixed effects) on students.

As a result, little is now about the relative size, and even the existence of endogenous and exogenous peer effects. However, this question is crucial because the presence of large endogenous effects implies the existence of a social multiplier : a shock to an individual, such as a policy, is amplified by the existence of social interactions. Therefore, the real effect of a policy can be much bigger than expected, in a way similar to the general equilibrium effects describe in the treatment effect literature (Miguel and Kremer 2004 [129]).

The only method enabling so far to consistently estimate all the parameters of the model is the one developed by Lee (2007) [117], and extended by Davezies et al. (2009) [52], which is used in this paper. It uses variation in class size in order to estimate all the parameters of a variation of the classical linear in mean model. Indeed, inside a given group, each individual is facing a different peer group, since this one does not include his or her own characteristics. This gives an opportunity to estimate consistently the parameters associated with the endogenous peer effects, while controlling for correlated effects and sorting across groups by using group fixed effects (which are no longer collinear with the measure of peer effects). Lee proposes a maximum likelihood estimator as well as a two stage least square estimator exploiting this opportunity. However, the convergence of these estimators is quite slow, and requires a large number of groups. As a result, it has been little exploited so far, with the exception of few papers exploiting the network structure recorded in some database to estimate peer effects in a closely related fashion (Lin, 2010 [120], Bramoullé et al., 2009 [31]), and with a recent paper by Boucher et al. (2014) [29] which uses this method but with a quite small database which prevents the authors to find precise enough estimates.

In this paper, I use a very large French database covering all French public and private school students which enables to apply this method with some credibility. The data-

base contains more than three millions students and one hundred and fifty thousands classrooms, which enables to obtain the convergence of Lee's maximum likelihood estimator, and to estimates exogenous and endogenous peer effects quite precisely.

Contrary to the previous literature based on network, I find no significant endogenous effect, and I can rule out medium to large endogenous peer effects thanks to the precision of my estimates. As for exogenous effects, I find quite large and significant effects only for some characteristics, particularly the share of repeaters in a classroom. This highlights the importance of fully estimating all the parameter of the structural model in order to assess correctly the size of endogenous and exogenous peer effects. It also confirms the previous findings on the small size of peer effects in an educational setting, on the existence of peer effects coming from disruptive and very weak students, and on the absence of a social multiplier as far as education is concerned.

The paper is organized as follows. The first part describes the data. The second explains the identification strategy. The third part exposes the results. The last part concludes.

## 5.2 Data and institutional context

I use administrative data from the French ministry of education at the level of the junior high school. I can use only data on students enrolled in the final year of junior high school, since I can then have access to their test score in a national exam. Each student can be matched at a high school, and inside the high school, to a classroom identifier. For most students, I have access to their test score at "*Brevet des collèges*", a national end of junior high school exam. This exam covers three different subjects : math, French, and history and geography. It is corrected anonymously at the regional level, which means that there is no scope for cheating or manipulating data as it has been reported in other context (Jacob and Levitt 2003 [102]).

The data includes students studying both in the public and the subsidized private sector. In practice, there is little difference between the two, since private schools who receive subsidies from the state have to follow a national curriculum, and to recruit their teachers through a national examinations organized by the state. Their staff is also paid by the state according to a national grid.

There is no common identifier between the database gathering students' characteristics and the one registering students results in the exam. As a result, I have to rely on probabilistic matching, based on students characteristics (namely birth date and school) which are present in the two datasets. I am able to match about 94,5% of the students to their results.

The data includes various information on students, including an indicator of their social background, their familial situation, the type of course load they have, and if they have already repeated a grade.

I had access to data on five cohorts, from 2006 to 2010. As a result, the size of the database is very large, and it includes about 3.5 millions students (3.2 millions if I keep only those who have test score and who can be matched to a classroom), and about 150,000 different classrooms.

Some descriptive statistics are presented below.

---

Table 1 : Descriptive statistics for the public and private French students enrolled in French 9th grade in 2006-2010

---

Variables	(N=3,627,298)
Percent Professional	0.169
Percent Semi-Professional	0.148
Percent Independent workers	0.084
Percent Agricultural	0.019
Percent White collar	0.166
Percent Blue collar	0.277
Percent Retired	0.018
Percent Inactive	0.094
Percent foreigners	0.036
Percent mother is responsible	0.200
Percent no other adult responsible	0.169
Percent female	0.505
Percent repeater	0.331
Average late years	0.382
Percent Scholarship	0.151

---

There is a large sorting across students, based on both characteristics and academic achievement. I regress the mean of the group (here, defined as a classroom) on the

results in math and French and find quite large and very significant coefficients, 0.21 for French, and 0.25 for math. It shows that students are strongly sorted into classrooms based on their academic results. This is also true for other characteristics which are correlated with scores, in particular socio-economic status.

As a result, endogenous sorting is likely to create bias if I naively try to estimate peer effects using a simple regression of academic results on group mean characteristics and other control variables.

## 5.3 Identification strategy

### 5.3.1 The econometric model

I estimate peer effects through a variation of the linear in mean model. Instead of measuring the peer group characteristics by the mean of the group, I measure it by the mean of all other students of the group. As a result, each students have a specific peer group, which allow to estimate a model with classroom fixed effects, since peer measure are no longer collinear with the classroom fixed effect. The full model is as follow :

$$y_{ri} = \lambda \left( \frac{1}{m_r - 1} \sum_{j=1, j \neq i}^{m_r} y_{rj} \right) + x_{ri} \beta_1 + \left( \frac{1}{m_r - 1} \sum_{j=1, j \neq i}^{m_r} x_{rj} \right) \beta_2 + \alpha_r + \epsilon_{ri} \quad (5.1)$$

The student  $i$  is enrolled in group  $r$ , which have a size  $m_r$ . His peer group includes all other students  $j$  of his group. His results depend on his peer group result mean  $\frac{1}{m_r - 1} \sum_{j=1, j \neq i}^{m_r} y_{rj}$ , his own exogenous characteristics  $x_{ri}$  and the mean of the exogenous characteristics of his peer group  $\frac{1}{m_r - 1} \sum_{j=1, j \neq i}^{m_r} x_{rj}$ . It also depends on a fixed effect  $\alpha_r$ , which capture both correlated effect (common factors affecting all the students of a group) and the common unobservable component across students of the same group (which emerge due to non random sorting into group).

In order to get rid of the fixed effects, it is possible to apply a panel data transformation :

$$(y_{ri} - \bar{y}_r) = -\lambda \frac{(y_{ri} - \bar{y}_r)}{m_r - 1} + (x_{ri} - \bar{x}_r) \beta_1 - (x_{ri} - \bar{x}_r) \left( \frac{\beta_2}{m_r - 1} \right) + (\epsilon_{ri} - \bar{\epsilon}_r) \quad (5.2)$$

The reduced form is :

$$(y_{ri} - \bar{y}_r) = (x_{ri} - \bar{x}_r) \left( \frac{(m_r - 1)\beta_1}{m_r - 1 + \lambda} \right) - (x_{ri} - \bar{x}_r) \left( \frac{\beta_2}{m_r - 1 + \lambda} \right) + \left( \frac{m_r - 1}{m_r - 1 + \lambda} \right) (\epsilon_{rj} - \bar{\epsilon}_r) \quad (5.3)$$

This model is identified only if there are different group sizes in the data. Otherwise, it is not possible to estimate separately  $\beta_1$ ,  $\beta_2$  and  $\lambda$ , since they are part of a ratio multiplying  $(x_{ri} - \bar{x}_r)$ . At least three different group sizes are needed, since there are three different parameters to estimate.

Lee (2007) [117] proposes two estimators for this model. The first one is based on conditional maximum likelihood, while the second rely on two stage least square, the instrument for  $\frac{(y_{ri} - \bar{y}_r)}{m_r - 1}$  being  $(x_{ri} - \bar{x}_r) \left( \frac{(m_r + 1)\beta_2}{m_r - 1 + \lambda} \right) - (x_{ri} - \bar{x}_r) \left( \frac{\beta_2}{m_r - 1 + \lambda} \right)$ . Lee investigates the properties of these estimators, and find through Monte-Carlo simulation that, while consistent, they are biased in small sample, and that their speed of convergence is very slow. In particular, the 2SLS estimator performs poorly. As a result, I use only the maximum likelihood estimator, for which I can obtain a reasonable convergence thanks to the size of my data.

Lee calculate the likelihood function for the model exposed above, assuming normal error :

$$L_{w,n}(\theta) = \prod_{r=1}^R \frac{\sqrt{m_r}}{(2\pi)^{(m_r-1)/2}} \left( \frac{m_r - 1 + \lambda}{\sigma(m_r - 1)} \right)^{(m_r-1)} \exp \left( -\frac{1}{2\sigma^2} \left( \frac{(m_r - 1 + \lambda)}{m_r - 1} Y_r^* - Z_r^* \delta_m \right)' \left( \frac{(m_r - 1 + \lambda)}{m_r - 1} Y_r^* - Z_r^* \delta_m \right) \right) \quad (5.4)$$

The star indicates that the variables are deviated from their means. As for  $Z$ , it is the vector made of  $(m_r + 1)(x_{ri} - \bar{x}_r)$  and  $(x_{ri} - \bar{x}_r)$ .  $\delta$  is the vector of the parameters  $[\beta_1, \beta_2]$ .

In the first appendix, I give more details on the procedure used to maximize this likelihood.

### 5.3.2 Missing data problem

Given that I cannot match all students to their results, I have a missing data problem. Missing data is important here, because I need to precisely estimate the peer group for each student, as well as the size of his or her class.

Davezies et al. (2009) [52] investigate this problem and show that, provided data are missing at random, it is possible to estimate the model using only the average of the

students matched to their results. In this case, the effect of the missing students on the other one is captured by the classroom fixed effect. As a result, the equation to be estimated become :

$$y_{ri} = \lambda \left( \frac{1}{m_r - 1} \sum_{j'=1, j' \neq i}^{m_r} y_{rj'} \right) + x_{ri} \beta_1 + \left( \frac{1}{m_r - 1} \sum_{j'=1, j' \neq i}^{m_r} x_{rj'} \right) \beta_2 + \alpha_r + \epsilon_{ri} \quad (5.5)$$

I denote by  $j'$  the other member of the group  $r$  for which there is no missing information. The new fixed effect embodies both the original classroom fixed effect as well as the effect of the missing students on the others. It is given by the equation below :

$$\tilde{\alpha}_r = \alpha_r + \lambda \left( \frac{1}{m_r - 1} \sum_{j^*=1, j^* \neq i}^{m_r} y_{rj^*} \right) + \left( \frac{1}{m_r - 1} \sum_{j^*=1, j^* \neq i}^{m_r} x_{rj^*} \right) \beta_2 \quad (5.6)$$

where I denote by  $j^*$  the students of the group  $r$  whose characteristics are missing.

However, in some case, missing data are clearly not arising at random. For instance, some students choose not to take the test because of their very weak results during the school year, which impede their ability to pass the exam. Indeed, the degree is awarded based on a combination of continuous assessment and a final exam. As a result, non test takers are very likely to be much weaker than the other students. Therefore, dropping this student is going to create a bias, since the peer effects will be measured only on a subset of stronger students. However, many studies point to the fact that peer effects are mostly driven by very weak students. As a result, this omission is problematic. I keep only classrooms where the number of students missing is not too high, below 3. This enables to use 80% of the sample.

In a specific case, I can use imputation methods. Indeed, for some students, while I do not have access on test score, I know the continuous assessment of the student. In this case, I replace test score by continuous assessment grades, taking into account the easiness of the grading. To do so, I remove the difference between the mean of the class in the test and in the continuous assessment, and I divide by the ratio of the variance of the two outcomes. The advantage of this solution is that the continuous assessment is impacted by the same peer as the test scores. As a result, it is an alternative measure of the same outcome.

## 5.4 Estimation and results

I present the result in this section. We perform the estimation on different samples, from the one where there is no missing data until the one where I include all classroom where less than three students are missing. I perform the estimation for two subjects, math and French.

The first column presents the estimate based on the sample including only classrooms with no missing students. The second column presents the estimates based on the sample including also the students whose grades can be inferred from their continuous assessment. The third column adds to the sample the students whose marks cannot be attributed unambiguously, because they are exactly similar in term of observables to another student in their cohort. Since this happens completely at random, they can safely be treated as randomly missing variable, and the inclusion of the sample should not lead to any bias. The last column includes all students in classrooms with at most two missing students.

I first present the estimated results for math. The endogenous effect is small, between 0 and 2 %. However, its standard error is large, up to 0.04 even when I use a sample with 2.6 millions students and close to 120,000 classrooms. This reflects the poor properties of the estimator. However, this is enough to rule out the existence of large endogenous peer effects. The last specification uses the largest sample. However, this sample includes classrooms with students not going to the test. This is likely to create a bias. Indeed, estimates for the exogenous peer effects appear to be extremely large and implausible. They are often larger than the scale of the test.

Across all specification, the effect of having more repeaters among one's peers is negative and significant. The magnitude of the effect is large, close to one standard deviation of test score. However, this correspond to the hypothetic case where all the member of the peer group have at least repeated one year in the past. Given that there is about 25 students in average in a classroom, the impact of an additional student being one year late is about 3 to 4 % of a standard deviation.

Similarly, students having retired parents have a very negative effect on other students. This is not so surprising given that people retiring early often have a low socio-economic status, since early retirement is the most frequent among workers having low paid and difficult jobs, and people victim of group layoff.

As for students with farmers parent, they have a very positive effect on the other. However, the effect is probably overestimated, given that the number of such students is small, and that the estimate diminished when the sample grows. The effect could be spurious and due to a bias in small sample.

The sign and magnitude of some coefficients change across specification. In general, all these estimates should be interpreted with some care. Their variance is large. Moreover, while consistent, the maximum likelihood estimator is likely to be biased in finite sample. In practice, Monte-Carlo simulation often found a bias (Lee 2007 [117]). As a result, it may be that the significant and large results found in this article are in fact due to the biasedness of the estimator. The fact that coefficients vary quite a lot, and that they sometimes change in sign reinforce this interpretation.

Moreover, the standard errors should be taken with care. Indeed, they are computed using the asymptotic variance calculated from the Fischer information matrix. However, this formula is correct only asymptotically. It may be that the variance of the

coefficient are different, and larger, in small sample.

I plan to use additional year of data in the future in order to improve the precision of the estimates.

However, some robust findings emerge. The coefficient of repeaters is stable across all samples, which indicate that it is fairly reliable. Similarly, the endogenous effect does not change much.

---

Table 2 : Peer effects in Math

---

Variables	No missing data N=488,552 M=20,690	continuous assessment N=678,408 M=28,812	no single match N=973,251 M=41,792	up to two non-random missing students N=2,680,532 M=117,333
Endogenous effect	0.0051 (0.191)	0.0105 (0.169)	0.0126 (0.071)	0.0217 (0.040)
repeater	-7.75*** (2.04)	-7.58*** (3.02)	-9.14*** (1.13)	-107.28*** (9.69)
female	0.51 (1.16)	0.84 (1.01)	-6.94*** (0.98)	-1.61 (1.36)
Scholarship	-2.25 (1.73)	-2.9** (1.44)	-3.76*** (1.14)	-32.64*** (3.79)
Professional	2.25 (2.69)	1.06 (2.55)	0.50 (2.16)	17.42*** (3.52)
Semi-Professional	3.75 (2.42)	4.37** (2.19)	1.45 (1.81)	32.10*** (3.15)
Independent workers	2.23 (1.81)	2.73* (1.57)	-2.41* (1.30)	18.16*** (2.07)
Agricultural	17.61*** (4.75)	13.82*** (3.74)	7.11** (3.16)	49.78*** (5.43)
White collar	6.50 (2.91)	8.75 (2.65)	-1.14 (1.70)	33.42*** (3.54)
Retired	-1.61** (1.51)	-1.57*** (1.35)	-11.53*** (1.23)	30.80*** (2.76)
Inactive	6.82*** (2.36)	6.24*** (1.98)	-1.68 (1.37)	34.24*** (3.07)

---

I then present the results for French. They are quite similar to the one for math. The endogenous effect is quite small and never significant, and its variance is large. However, it is in general a bit higher than the effect in math. As for the exogenous effect, I find again a large, significant and negative effect of studying with peers having repeated a class before, across all sample, in similar fashion to the math estimates. The other coefficients vary a lot across samples. They sometimes change signs, as for female, inactive, white collar or independent workers. As a result, little confidence should be given to these estimates, as in math, except for the coefficient for repeaters, which is quite stable and seems credible.

Table 3 : Peer effects in French

---

Variables	No missing data N=506,193 M=21,441	continuous assessment N=680,421 M=28,899	no single match N=964,000 M=41,899	up to two non-random missing students N=2,842,025 M=117,423
Endogenous effect	0.025 (0.169)	0.058 (0.156)	0.034 (0.035)	-0.004 (0.043)
repeater	-5.22*** (1.23)	-5.52*** (1.17)	-12.40*** (0.81)	-104.62*** (9.98)
female	-0.10 (0.79)	0.61 (0.78)	-8.48*** (0.70)	4.64*** (1.36)
Scholarship	-1.58 (1.22)	-2.14* (1.11)	-7.22*** (0.90)	-18.93*** (2.76)
Professional	2.11 (1.98)	1.61 (2.01)	12.61*** (1.76)	4.05 (3.52)
Semi-Professional	4.51** (1.91)	4.79*** (1.85)	1.83 (1.37)	14.11*** (2.70)
Independent workers	2.68* (1.38)	2.24* (1.25)	-6.72*** (1.00)	0.60 (1.72)
Agricultural	13.39*** (3.24)	13.09*** (2.98)	1.53 (2.32)	35.70*** (5.89)
White collar	5.47*** (2.10)	6.92*** (2.00)	-14.69*** (1.43)	25.13*** (3.63)
Retired	-1.37 (1.08)	-1.65 (1.08)	-11.53*** (0.80)	-6.18*** (1.98)
Inactive	3.29** (1.46)	2.95** (1.35)	-6.40*** (1.04)	5.61*** (1.98)

---

## 5.5 Conclusion

I estimate endogenous and exogenous peer effects using group size variation. This is interesting since it gives a direct estimate of the endogenous effect, which is in general estimated by using lagged test scores, or instrumental variables. The endogenous effect is not significant from zero, and I can reject medium to large effect. As a result, we can cast a doubt on the existence of a social multiplier in schools.

Some exogenous effects are also significant and quite large, in particular the effect of studying with repeaters, which is stable across all sample used.

However, the variance of the estimates is large, and some are probably biased, given the poor properties of the estimator. I need to use a larger dataset to alleviate this problem.

## 5.6 Appendix 1 : Detail of the procedure to maximize the log-likelihood

The log likelihood function concentrated at  $\lambda$  is

$$\begin{aligned} \ln L(\lambda) = c_1 + \sum_{r=1}^R (m_r - 1) \ln(m_r - 1 + \lambda) \\ - \frac{R(m-1)}{2}. \\ \ln\left(\frac{1}{R_m - 1} \left(\sum_{r=1}^R \left(\frac{m_r - 1 + \lambda}{m_r - 1}\right)^2 Y_r^{*'} Y_r^* - \sum_{r=1}^R \left(\frac{m_r - 1 + \lambda}{m_r - 1}\right)^2 Y_r^{*'} Z_r^* \left(\sum_{r=1}^R Z_r^{*'} Z_r^*\right)^{-1} \sum_{r=1}^R \left(\frac{m_r - 1 + \lambda}{m_r - 1}\right)^2 Z_r^{*'} Y_r^*\right)\right) \end{aligned}$$

The part in the log in the right hand side can be rewritten linearly, and decomposed in several elements. Let's denote :

$$A = \sum_{r=1}^R (m_r - 1) \ln(m_r - 1 + \lambda) \quad (5.8)$$

$$\begin{aligned} B = \sum_{r=1}^R \left(\frac{m_r - 1 + \lambda}{m_r - 1}\right)^2 Y_r^{*'} Y_r^* = \lambda^2 \sum_{r=1}^R \left(\frac{Y_r^{*'} Y_r^*}{(m_r - 1)^2}\right) \\ + 2\lambda \sum_{r=1}^R \left(\frac{Y_r^{*'} Y_r^*}{(m_r - 1)}\right) + \sum_{r=1}^R Y_r^{*'} Y_r^* \quad (5.9) \end{aligned}$$

$$\begin{aligned} C = \lambda^2 \sum_{r=1}^R \left(\frac{1}{m_r - 1}\right) Y_r^{*'} Z_r^* \left(\sum_{r=1}^R Z_r^{*'} Z_r^*\right)^{-1} \sum_{r=1}^R \left(\frac{1}{m_r - 1}\right) Z_r^{*'} Y_r^* \\ + 2\lambda \sum_{r=1}^R \left(\frac{1}{m_r - 1}\right) Y_r^{*'} Z_r^* \left(\sum_{r=1}^R Z_r^{*'} Z_r^*\right)^{-1} \sum_{r=1}^R Z_r^{*'} Y_r^* \\ + \sum_{r=1}^R Y_r^{*'} Z_r^* \left(\sum_{r=1}^R Z_r^{*'} Z_r^*\right)^{-1} \sum_{r=1}^R Z_r^{*'} Y_r^* \quad (5.10) \end{aligned}$$

Using these notations, the log-likelihood function becomes :

$$\ln L(\lambda) = c_1 + A - \frac{R(m-1)}{2} \ln\left(\frac{1}{R_m - 1} (B - C)\right) \quad (5.11)$$

It is now easy to maximize this likelihood with respect to  $\lambda$ . The first order condition is :

$$\frac{\partial \ln L(\lambda)}{\partial \lambda} = A' - \frac{R(m-1)}{2} \frac{(B' - C')}{(B - C)} = 0 \quad (5.12)$$

with :

$$A' = \sum_{r=1}^R \frac{(m_r - 1)}{(m_r - 1 + \lambda)} \quad (5.13)$$

$$B' = 2\lambda \sum_{r=1}^R \left( \frac{Y_r^{*'} Y_r^*}{(m_r - 1)^2} \right) + 2 \sum_{r=1}^R \left( \frac{Y_r^{*'} Y_r^*}{(m_r - 1)} \right) \quad (5.14)$$

$$\begin{aligned} C' = 2\lambda \sum_{r=1}^R \left( \frac{1}{m_r - 1} \right) Y_r^{*'} Z_r^* \left( \sum_{r=1}^R Z_r^{*'} Z_r^* \right)^{-1} \sum_{r=1}^R \left( \frac{1}{m_r - 1} \right) Z_r^{*'} Y_r^* \\ + 2 \sum_{r=1}^R \left( \frac{1}{m_r - 1} \right) Y_r^{*'} Z_r^* \left( \sum_{r=1}^R Z_r^{*'} Z_r^* \right)^{-1} \sum_{r=1}^R Z_r^{*'} Y_r^* \quad (5.15) \end{aligned}$$

As a result, I obtain a non linear equation in  $\lambda$ , which can be solved by using classical methods.



## Quatrième partie

### Le marché du travail des professeurs



## Chapitre 6

# La gestion des ressources humaines dans le secteur éducatif

## 6.1 introduction

Pendant longtemps, les enseignants n'ont pas été un enjeu dans le débat éducatif en France. Les discussions ont plutôt porté sur les programmes (avec, par exemple, la discussion des nouveaux programmes introduits par Luc Ferry), les pédagogies (avec notamment le fameux débat sur la méthode globale), la carte scolaire (avec la réforme Sarkozy de 2007), ou encore l'orientation (avec le débat sur les mérites du collège unique).

Cependant, une série de réformes et de rapports officiels ont réintroduit les enseignants dans la discussion sur l'éducation. Tout d'abord, la réforme Sarkozy supprimant les IUFM et introduisant l'obligation d'obtenir un master pour pouvoir passer les concours de l'enseignement a mis au premier plan la question de la formation des enseignants. De même, la généralisation des heures supplémentaires a introduit de manière indirecte la question du statut des enseignants, question qui a de nouveau été abordée par un récent rapport de la cour des comptes appelant à une réforme en profondeur de ce statut. En effet, le système de gestion des enseignants en France apparaît comme très rigide (Cour des comptes, 2013 [54]). Les enseignants sont affectés aux écoles en fonction d'une règle administrative fondée sur l'ancienneté et sur leurs vœux, sans que les caractéristiques des écoles ou encore l'avis du directeur d'établissement ne soient pris en compte. Il existe par ailleurs très peu de différenciation et de spécialisation entre enseignants, qui ne sont officiellement pas sous la tutelle hiérarchique du directeur d'établissement. Ce dernier n'a qu'un rôle de coordination de l'équipe pédagogique. Les enseignants sont évalués de manière très irrégulière par un corps d'inspecteurs, et cette évaluation a relativement peu d'influence sur le déroulement de leur carrière, où la rémunération au mérite reste faible par rapport à l'ancienneté. Les enseignants sont donc pour l'essentiel laissés sans supervision, sans que l'administration ait une vraie politique de ressource humaine. C'est notamment cette lacune qui suscite aujourd'hui le débat.

Progressivement, deux grandes positions ont émergé. Une position que l'on pourrait qualifier d'attentiste, ou encore de sceptique, consiste à vouloir préserver autant que possible l'autonomie pédagogique des enseignants, et à se méfier de toute tentative pour renforcer leur tutelle hiérarchique, leur évaluation, ou encore la part liée au mérite de leur rémunération. Le principal argument de cette approche est la peur de la mise en place d'un système arbitraire et clientéliste où les enseignants seraient placés sous le contrôle étroit de bureaucrates n'ayant pas les compétences nécessaires pour les évaluer. En effet, l'enseignant est seul face à sa classe, et il est difficile de mesurer son impact sur le long terme. Dès lors, l'évaluation serait illusoire, et ne serait qu'un moyen pour asseoir le pouvoir d'une administration centralisée. La seule politique éducative souhaitable serait donc d'augmenter la rémunération des enseignants afin d'attirer un vivier de candidats ayant un bon niveau académique, et d'augmenter leur nombre dans le but d'améliorer les conditions d'enseignement.

A l'inverse, une position que l'on pourrait qualifier d'activiste promeut une gestion

plus active des ressources humaines de l'éducation nationale. Elle affiche son scepticisme envers l'efficacité des moyens supplémentaires, soulignant que les résultats sont restés stables malgré la hausse du nombre d'enseignants par élève. Elle souligne l'importance d'attirer des enseignants plus motivés et de meilleure qualité en augmentant leur rémunération, quitte à réduire leur nombre pour financer la hausse de leur traitement et à accroître leur rémunération grâce à des heures supplémentaires. Enfin, elle valorise l'autonomie des établissements, et accorde un rôle central au proviseur, qui doit pouvoir constituer librement son équipe pédagogique, et jouer un plus grand rôle dans l'évaluation des enseignants, voire dans l'attribution de primes au mérite.

On le voit, ces deux conceptions sont à l'opposé l'une de l'autre. D'une certaine manière, elles reflètent aussi des différences d'idéologie. La vision sceptique pourrait être vue comme proche d'un point de vue politique « de gauche », notamment en raison de l'importance donnée à l'égalité de traitement entre enseignants, la peur du contrôle hiérarchique vu comme liberticide et arbitraire, et sa volonté de garantir un service public homogène sur le territoire en réduisant l'autonomie des établissements et en traitant les enseignants de manière identique. A l'inverse, la vision activiste pourrait être vue comme plus proche d'un point de vue politique « de droite », notamment en raison de son insistance sur l'initiative personnelle, l'autonomie, les différences de capacité entre individus, et la nécessité de récompenser les meilleurs enseignants et de donner plus d'incitations. Il n'est d'ailleurs pas surprenant qu'une partie des éléments qui la composent ait été mise en œuvre durant le mandat de Nicolas Sarkozy, ni que d'autres aient fait partie du programme de l'UMP.

Cependant, il faut se garder de réduire ce débat à un simple conflit idéologique. Tout d'abord, de nombreuses personnes ont une position nettement plus nuancée, et empruntent simultanément des éléments aux deux visions exposées ci-dessus. Ainsi, les conclusions du rapport de la Cour des comptes ont-elles été bien accueillies par la plupart des tendances politiques, l'idée que le statut des enseignants en France est inadapté faisant consensus. Surtout, le désaccord entre ces deux visions repose foncièrement sur des éléments empiriques. Est-il possible de mesurer la contribution d'enseignants individuels à la réussite des élèves ? Peut-on mettre en place des dispositifs incitatifs efficaces tout en évitant les effets pervers ? Les chefs d'établissements ont-ils une idée juste de la valeur des enseignants qu'ils supervisent ? Ces questions, parmi d'autres, sont au cœur du débat entre les deux conceptions que nous avons évoquées, et sont également des questions factuelles susceptibles d'être résolues par des études empiriques bien menées. Pendant longtemps, néanmoins, de telles études n'étaient pas réalisables dans la mesure où il était difficile de disposer de larges bases de données appariant simultanément les enseignants et leurs caractéristiques, celles de leurs élèves, et les résultats académiques de ces derniers. Cependant, ces dix dernières années, le développement de procédures d'évaluation des écoles aux Etats-Unis, notamment en raison de la loi « No child left behind », ont abouti à la mise à disposition des chercheurs de bases de données de

ce type. Cela a permis la production de nombreuses études empiriques permettant de mieux comprendre ce qui rend les enseignants plus ou moins efficaces, ainsi que les effets de différentes politiques de ressources humaines. Le but de cet article est donc de passer en revue cette abondante littérature empirique afin d'éclairer le débat sur la gestion des enseignants en France.

On s'interrogera tout d'abord sur l'existence de différences d'efficacité entre enseignants. Elles conditionnent en grande partie l'intérêt d'une politique active de ressource humaine. On s'interrogera ensuite sur les facteurs observables pouvant influencer l'efficacité des enseignants, et susceptibles d'éclairer les procédures de recrutement. On essaiera alors de faire le point sur l'impact de la formation des enseignants, qui est le principal moyen par lequel le système éducatif peut affecter leur efficacité après leur recrutement. On s'intéressera ensuite à l'efficacité des différents systèmes d'incitation et de rémunération au mérite ayant été expérimentés tant aux Etats-Unis que dans d'autres pays de l'OCDE. On discutera enfin la répartition des enseignants au sein du système éducatif.

## 6.2 Les différences d'efficacité entre enseignants

Tous les enseignants se valent-ils ? La question pourrait prêter à sourire, tant chacun a des souvenirs d'enseignants atrocement médiocres, ou au contraire d'enseignants charismatiques, capables de passionner toute une classe. Cependant, il est possible que les enseignants n'aient qu'un impact négligeable sur les différences de réussite scolaire par rapport à des facteurs tels que les effets de pair ou bien l'origine sociale des étudiants. Même si les différences perçues sont importantes, il est possible qu'elles ne pèsent que peu face à ces facteurs moins spectaculaires, mais peut-être plus influents. De fait, la littérature essayant de mesurer les facteurs de réussite des élèves s'est pendant longtemps concentrée sur des éléments tels que la taille des classes, les effets de pair, l'appartenance ethnique, ou encore l'origine sociale.

Néanmoins, depuis le début des années 2000, une littérature croissante s'est attachée à quantifier les variations en termes d'efficacité des enseignants, et à essayer de mesurer précisément la contribution d'un enseignant en particulier au progrès des élèves. La première étude robuste sur le sujet (Hanushek, Kain et Rivkin, 2005 [138]) a mis ainsi en évidence un écart-type de l'effet professeur non négligeable, de l'ordre de 10-15 % d'un écart-type de score par année de scolarité. Si cet effet semble faible à première vue, il faut garder à l'esprit que l'éducation est un processus cumulatif, et qu'un élève se voit assigner plusieurs enseignants au cours de sa scolarité. Par conséquent, se voir assigner une série de bons ou de médiocres professeurs peut avoir un effet très important.

Par la suite, la littérature sur l'effet professeur a, avant tout, essayé de proposer des mesures de valeur ajoutée fondées sur le calcul d'un effet fixe enseignant. Il s'agit en

substance de mesurer la moyenne des performances scolaires des élèves d'un enseignant donné, après avoir contrôlé les caractéristiques de ces élèves. Un problème majeur auquel doit faire face cette littérature est l'allocation non aléatoire des enseignants aux classes. En effet, il est possible que les bons professeurs soient systématiquement assignés aux meilleurs élèves, ou inversement, ce qui mènerait à une sous- (sur)estimation de l'écart-type de l'effet professeur. Par ailleurs, même si les enseignants sont assignés aux élèves de manière aléatoire conditionnellement aux observables, l'écart-type de l'effet professeur risque d'être surestimé dans la mesure où, par chance, des professeurs ayant un niveau d'efficacité identique peuvent se trouver, pour certains, avec une « bonne classe », et pour d'autres avec une « mauvaise » classe, ce qui conduit à surévaluer la différence d'efficacité entre ces enseignants. Une partie des différences de résultats observées entre enseignants serait ainsi due au hasard plutôt qu'à des différences d'efficacité réelles.

Par conséquent, une série d'articles (Rockoff, 2004 [139], Kane et Staiger, 2008 [], Aaronson et al., 2007 [1]) a tenté de résoudre ces deux problèmes à l'aide de techniques économétriques plus robustes. Tout d'abord, ces chercheurs ont essayé de contrôler du mieux possible le niveau des élèves en utilisant leurs résultats scolaires antérieurs, ce qui permet de limiter l'ampleur d'un biais lié à la répartition des classes et des professeurs, s'il existe. Ainsi, en utilisant de riches bases de données sur les étudiants, il est possible d'expliquer jusqu'à 80 % de la variance de leurs résultats. L'ampleur des biais possibles devient donc relativement réduite.

Par ailleurs, l'utilisation des estimateurs de réductions s'est généralisée. Ceux-ci accordent plus ou moins de poids aux effets professeurs en fonction de la précision avec laquelle ils sont estimés. Ainsi, un enseignant dont les résultats varient fortement d'une année sur l'autre, et dont les élèves ont eux-mêmes des résultats très hétérogènes au sein d'une même classe, se verra attribuer un poids très faible dans le calcul de la variance totale. A l'inverse, un enseignant qui, de manière répétée, a des résultats identiques et dont les élèves ont des résultats homogènes, se verra attribuer un poids important. Cela permet donc de prendre en compte l'impact de la chance par rapport à une mesure de l'effet professeur où tous les professeurs sont traités de manière similaire dans le calcul de la variance. Ces études ont trouvé des résultats proches de ceux de Hanushek, Kain et Rivkin, soit environ 10-15 % d'un écart type pour les enseignants en mathématiques, et 5-10 % pour les enseignants en français. L'existence d'une variation significative de qualité entre les enseignants semble donc confirmée.

Cependant, comme le fait remarquer Jesse Rothstein (2010) [148], même en contrôlant les résultats scolaires passés, il est toujours possible que des biais existent sur les inobservables, comme le montrent plusieurs tests statistiques qu'il met en œuvre pour montrer les limites des mesures de valeurs ajoutées. Par ailleurs, comme le font remarquer d'autres chercheurs, un accroissement des scores peut n'être que temporaire, et n'avoir aucun lien avec des résultats de long terme tels que les salaires ou l'obtention d'un diplôme universitaire. Carrell et West (2010) [43] trouvent ainsi, en observant une

université américaine, que les enseignants qui ont un impact négatif sur les résultats de fin d'année ont un effet positif sur les résultats des années suivantes, et réciproquement. Bien que difficilement généralisable en raison du contexte très spécifique où elle est menée, cette étude met cependant en garde contre une interprétation trop hâtive des mesures de valeur ajoutée comme capacité à faire progresser durablement les élèves. Par ailleurs, les mesures de valeur ajoutée n'évaluent que la capacité à faire augmenter les compétences cognitives, et ne disent rien sur l'efficacité des enseignants quant à l'augmentation des compétences non cognitives (telles que la capacité à respecter les règles, à faire des efforts, et ainsi de suite). Ainsi, Kirabo Jackson (2013) [98] montre par exemple que certains enseignants apparemment peu efficaces se révèlent très capables si on évalue leur habileté à faire augmenter ces compétences non cognitives.

Des travaux récents ont néanmoins permis de répondre à une partie de ces critiques, et de valider une grande partie des résultats antérieurs. Ainsi, Kane et Staiger (2008) [106] ont mis en œuvre une expérience contrôlée où un groupe d'enseignants était assigné aléatoirement à des classes, tandis qu'un autre était assigné de manière traditionnelle, en fonction des choix des proviseurs. L'étude ne trouve pas de trace de sélection sur les inobservables, les résultats étant sensiblement identiques entre les groupes. Par ailleurs, Chetty et al. (2013a [46], 2013b [48]) ont réussi à suivre sur une très longue période une série de cohortes d'enseignants et d'étudiants, et ont notamment pu obtenir des données sur les résultats des étudiants sur le marché du travail. Ils montrent ainsi que les élèves se voyant assigner un enseignant ayant une valeur ajoutée élevée réussissent significativement mieux sur le marché du travail, même des années plus tard. Se voir assigner au collège ou au primaire un enseignant ayant une valeur ajoutée supérieure à un écart-type par rapport à la moyenne est associé à une augmentation de salaire d'environ 1 %. Ils trouvent également des effets positifs sur la qualité du quartier habité, la probabilité d'éviter une grossesse précoce, et la probabilité de fréquenter une institution universitaire. Se voir assigner un professeur de bonne qualité, telle qu'elle est mesurée par les économistes, est donc bien associé à une amélioration de sa situation sur le long terme.

Chetty et al. proposent également des tests de robustesse fondés sur les changements d'enseignants entre écoles. En effet, quand un nouvel enseignant arrive dans une école, il fait changer le niveau moyen des étudiants de l'établissement quelle que soit la classe à laquelle il est assigné, supprimant le problème du biais lié au tri entre enseignants et étudiants. Comparer la réussite moyenne des étudiants de l'école avant et après son arrivée permet donc d'évaluer dans quelle mesure la valeur ajoutée estimée est sans biais. Chetty et al. trouvent un biais minime, validant la méthodologie utilisée jusque-là. Dans la mesure où les modèles économétriques utilisés expliquent une très grande part de la variance des scores des élèves, une telle absence de biais n'est pas forcément surprenante.

Enfin, Chetty et al. tentent également de quantifier les gains économiques générés par

un enseignant plus efficace que la moyenne. Ils trouvent des résultats très élevés, égaux à plusieurs centaines de milliers de dollars par an pour les enseignants dans le haut de la distribution. Réalisée sur un échantillon de plusieurs millions d'élèves, portant sur une période d'une vingtaine d'années et avec des données très riches, l'étude de Chetty et al. permet donc d'apporter une conclusion définitive au débat. Par conséquent, la littérature scientifique a atteint un relatif consensus sur l'importance des enseignants. Leur efficacité varie de manière importante, et il s'agit d'un facteur crucial dans la réussite académique des élèves, ainsi que dans leur réussite future dans un ensemble de domaines économiques et sociaux. Cependant, comme on va le voir plus loin, cela ne fait pas pour autant des enseignants un facteur aisément manipulable pour améliorer les résultats scolaires.

### 6.3 Le recrutement et la formation des enseignants

Si les enseignants diffèrent en termes de capacité, il est naturel de vouloir sélectionner en priorité les meilleurs d'entre eux. Cependant, avant le recrutement, il n'est pas possible de mesurer ces capacités, puisque les aspirants enseignants n'ont aucune expérience. Dès lors, les recruteurs se concentrent sur des caractéristiques aisément observables, dont on suppose qu'elles sont corrélées à la capacité d'enseigner. Ainsi, la plupart des systèmes éducatifs accordent un poids important à des critères académiques, tels que les programmes universitaires suivis, les diplômes obtenus, et surtout, la capacité à réussir un examen de certification des connaissances (c'est la voie suivie par la France). Par ailleurs, dans les systèmes où les changements d'affectation ne se font pas de manière administrative, mais à travers un marché externe et décentralisé, et où les écoles ne peuvent observer directement les performances passées dans d'autres établissements, un poids important est généralement donné à l'expérience. Il est donc important de mesurer l'impact réel de ces caractéristiques sur la réussite des élèves, ce qui permettrait de mettre en place des pratiques de recrutement efficaces.

Les principales études sur le sujet aux Etats-Unis ont été réalisées par Clotfelter et al. (2006) [50] à partir de données sur la Caroline du Nord. La principale difficulté dans l'estimation de l'impact de ces caractéristiques est l'existence d'un tri entre élèves et professeurs, les professeurs les plus qualifiés se voyant en général assigner les meilleurs élèves. Comme on l'a vu précédemment, même si on peut contrôler les caractéristiques observables des élèves, il est toujours possible qu'un tri subsiste à partir des caractéristiques inobservables. Ceci pourrait introduire un biais important dans l'estimation. Dans le cadre de leur première étude (2006), fondée sur des données sur des élèves de CM2 (fifth grade), Clotfelter et al. essaient de supprimer ce problème en contrôlant de nombreuses caractéristiques des élèves, incluant leurs résultats scolaires passés, en introduisant des effets fixes pour chaque école, et en gardant dans leur échantillon uniquement les écoles où il ne semble pas y avoir de tri fondé sur les observables entre

les classes. Cette méthodologie ne résout pas le problème de sélection sur les inobservables, mais en réduit cependant fortement l'ampleur. La seule caractéristique ayant un effet est l'expérience de l'enseignant, et ce seulement pendant les deux premières années. Après deux années d'expérience, les enseignants ne gagnent plus en efficacité avec l'âge. Le niveau académique, mesuré par la sélectivité de l'institution universitaire fréquentée, par les résultats aux tests de certification (licensure test score, l'équivalent du CAPES), et la possession d'un titre universitaire post-licence n'a aucun impact sur la réussite des élèves.

Dans une étude ultérieure (2010) [51], Clotfelter et al. étudient l'impact des caractéristiques des professeurs sur des étudiants inscrits au lycée. La méthodologie utilisée dans ce cas est le contrôle de la qualité des élèves en utilisant des effets fixes élèves à travers les différentes matières. On regarde donc si les caractéristiques des enseignants dans une matière sont corrélées avec l'écart des résultats des élèves dans cette matière par rapport à leur moyenne générale. Ceci permet de tenir compte du niveau général de l'élève. De cette manière, la seule source de biais possible est qu'un professeur ayant un certain type de caractéristiques soit systématiquement affecté à une classe où en moyenne les élèves sont plus forts dans une discipline plutôt que dans une autre. Un tel phénomène pourrait apparaître si, par exemple, les élèves étaient systématiquement regroupés en fonction de leur capacité dans une discipline par rapport aux autres (création d'une classe forte en mathématiques, d'une autre classe forte en français, et ainsi de suite). En pratique, Clotfelter et al. ne repèrent pas un tel modèle sur les caractéristiques observables dans leurs données, ce qui rend peu probable l'existence d'un important tri sur les inobservables.

Les résultats sont assez différents de la précédente étude, puisque la plupart des caractéristiques étudiées ont un impact. L'expérience a un effet approximativement similaire au cas du primaire. Par ailleurs, le fait d'être certifié (au lieu d'être recruté par une voie alternative) est assez important (5 à 7 % d'un écart-type), de même que le fait d'être certifié par le National Board for Certification, organisme évaluant la qualité d'un enseignant à partir de l'observation répétée de son comportement en classe (5 %). Cependant, même si les caractéristiques académiques ont un impact, celui-ci est extrêmement faible. Ainsi, l'impact du test à l'examen de certification est de 0,7 % d'un écart-type pour un enseignant ayant un score supérieur d'un écart-type à la moyenne (soit un enseignant se trouvant dans les 15 % les meilleurs à l'examen). De même, l'impact du niveau de sélectivité de l'institution universitaire fréquentée est faible ou nul. Par ailleurs, le sexe et l'origine ethnique ont une influence importante, à hauteur de 6 % chacun, mais ne sont guère susceptibles d'être utilisés dans le recrutement pour des raisons de non-discrimination.

Ces résultats sont confirmés par Harris et Sass (2011) [82] à partir de données venant du système scolaire en Floride, et avec une méthodologie proche. Disposant de données détaillées sur les cours pris par les enseignants à l'université durant leurs études, ils ne trouvent pas d'impact d'un type de cours en particulier sur l'efficacité des enseignants,

quel que soit le niveau auquel ils enseignent. De manière similaire, ils ne trouvent aucun effet de la capacité académique des enseignants sur la réussite des élèves. On peut faire l'hypothèse que la certification a plus d'importance au lycée en raison du plus haut niveau académique exigé de l'enseignant. Cependant, dès lors que ce niveau minimal est atteint, les compétences académiques ne jouent plus de rôle important.

On le voit donc, les caractéristiques observables susceptibles d'être utilisées pour recruter les enseignants, et en particulier les performances académiques, n'ont guère d'impact sur la réussite des étudiants. La seule caractéristique importante est le fait de se voir assigner un enseignant débutant mais elle n'est guère manipulable par les pouvoirs publics. Par conséquent, le processus de recrutement ressemble à un exercice de sélection à l'aveugle où les écoles sont obligées de départager les candidats sur des bases arbitraires. Si le problème de l'absence d'information est difficilement soluble lors du recrutement, il est cependant susceptible d'être résolu dans les années suivantes. En effet, il est alors possible de disposer de mesures de l'efficacité des enseignants fondées sur les résultats de leurs élèves. Dès lors, on peut imaginer un dispositif de titularisation où les jeunes enseignants passeraient à travers une période probatoire durant plusieurs années, avant d'être confirmés dans leur poste dès lors que leurs résultats seraient jugés suffisamment bons. On pourrait également inciter les professeurs les plus médiocres à quitter l'enseignement en liant une part importante de la rémunération aux résultats. Cependant, cela suppose d'être capable de mesurer précisément l'efficacité des enseignants à leur poste de travail. Or, si les travaux passés en revue précédemment montrent bien l'existence d'une forte hétérogénéité entre enseignants, leurs estimations des effets professeurs individuels sont de faible qualité. En effet, ces études procèdent en deux étapes. Tout d'abord, elles estiment un effet professeur pour chaque enseignant. Il s'agit de regarder si l'enseignant fait progresser plus ou moins fortement ses élèves par rapport à la moyenne des autres professeurs (qui par convention est fixée à zéro). Ensuite, elles mesurent l'écart-type de ces effets professeurs (qui sont considérés comme étant une variable aléatoire). Comme on l'a vu précédemment, on peut considérer qu'en moyenne l'effet mesuré pour un enseignant donné est sans biais, comme le montrent de manière convaincante Chetty et al. (2011). Cependant, chacun de ces effets individuels est mesuré avec une marge d'erreur, qui peut être importante. En effet, il existe la plupart du temps un faible nombre d'observations par enseignant, ce qui fait exploser la variance des effets professeurs tels qu'ils sont mesurés. Le problème est exacerbé si les enseignants évalués n'ont que quelques années d'expérience.

Ce n'est pas une difficulté dès lors que l'on se concentre sur la mesure de la variance de l'effet professeur, puisque c'est principalement l'existence d'un biais qui compte, et que l'effet de la variance des effets professeurs individuels sur la mesure de la variance de l'ensemble des effets professeurs peut être neutralisée par l'utilisation d'estimateur de réduction, comme on l'a vu plus haut. En revanche, cela devient un obstacle majeur dès lors que l'on veut utiliser ces informations pour prendre des décisions de titularisation ou de paie au mérite, qui nécessitent notamment de comparer les enseignants les uns

par rapport aux autres. Dans ce cas, le risque d'erreur devient très important. Cependant, on pourrait considérer qu'en moyenne, les professeurs titularisés sont meilleurs que ceux qui ne le sont pas, et que dès lors il est rentable de mettre en place un tel dispositif. C'est l'approche proposée par Staiger et al. (2010) [156] qui proposent de recruter massivement de jeunes professeurs et de n'accorder la titularisation qu'à un nombre très réduit d'entre eux, de l'ordre d'un quart, voire moins. Dès lors que le taux de sélection est très élevé, même une mesure très bruitée des capacités des professeurs peut permettre d'améliorer de manière très sensible l'efficacité moyenne des professeurs.

Une telle approche suppose néanmoins que l'offre de travail pour le métier de professeur est très élevée, ce qui en général n'est pas le cas. En effet, dès lors qu'une part importante des jeunes professeurs n'est pas titularisée, il faut recruter chaque année un nombre d'enseignants largement supérieur aux besoins (puisque seuls certains enseignants auront vocation à rester dans le système scolaire sur le long terme). Staiger et al. objectent qu'utiliser un système de recrutement fondé sur une mesure sans biais de l'effet professeur permettrait d'élargir de manière considérable le réservoir des professeurs potentiels en réduisant fortement les prérequis nécessaires pour enseigner, prérequis qui en pratique ne sont que faiblement corrélés avec la capacité à faire progresser les élèves.

Cependant, même en assouplissant les barrières à l'entrée, l'augmentation nécessaire du nombre de candidats ne semble pas réalisable. En effet, environ 4 % de la force de travail exercent en tant qu'enseignant dans la plupart des pays de l'OCDE, ce qui représente de l'ordre de 8 à 10 % des diplômés du supérieur. Dans un système où la titularisation est quasi systématique, cela signifie que le système éducatif doit approximativement réussir à attirer 8 à 10 % des jeunes diplômés chaque année, ce qui n'est déjà pas négligeable et peut devenir problématique dans certains pays. Par conséquent, ne titulariser que 50 % des jeunes enseignants reviendrait à s'imposer l'obligation d'attirer chaque année de l'ordre de 15 à 20 % des nouveaux diplômés du supérieur dans les carrières de l'enseignement, ce qui est considérable. Or Staiger et al. sont en faveur de taux de titularisation encore plus faibles.

Par ailleurs, Rothstein (2012) [149] fait observer qu'un taux élevé de sélection dégrade considérablement l'attrait de la profession d'enseignant. En effet, l'expérience acquise dans l'enseignement n'est généralement pas transférable à d'autres professions, et ce d'autant moins que l'enseignant a été licencié, ce qui envoie un signal négatif au futur employeur. Attirer des candidats requerrait donc une hausse massive de leur rémunération, au moins lors des premières années, afin de compenser les risques considérables pris. Or, plus la période de titularisation est élevée, plus le montant de la compensation doit être élevé. A l'inverse, plus la période de titularisation est faible, plus la variance de l'effet professeur mesuré est forte.

Par conséquent, on voit que les problèmes liés à l'offre de travail des enseignants viennent sévèrement limiter les possibilités de mise en place d'un système de titularisa-

tion. L'existence d'un biais dans la mesure de l'effet professeur viendrait encore empirer la situation, en réduisant les gains d'efficacité à espérer d'une plus grande sélection des professeurs. Finalement, sous des paramètres plausibles, et en faisant l'hypothèse d'une absence de biais des mesures de qualité des professeurs, Rothstein trouve des coûts très élevés sous forme de hausses de salaire aux politiques visant à augmenter le taux de licenciement des professeurs les plus faibles, ou à mettre en place un système de bonus rémunérant la performance. Ainsi évalue-t-il à 30 % la hausse de la masse salariale nécessaire à la mise en place d'un taux de licenciement de 20 % en début de carrière. Dès lors que le mécanisme de licenciement et de bonus est biaisé, l'efficacité du système est réduite de manière considérable, alors que les coûts restent identiques. Par conséquent, le modèle proposé par Rothstein montre les limites d'une politique de titularisation débouchant sur un taux de licenciement important. Si elle peut donner des bénéfices réels, une telle politique a cependant des coûts qui deviennent rapidement exorbitants. Il est sans doute préférable de mettre en œuvre des politiques de titularisation et de bonus relativement mesurées.

Une autre piste serait d'améliorer les mesures de performance des enseignants, notamment en utilisant des observations directes du professeur dans sa classe plutôt que des mesures fondées sur des scores. Cependant, les expériences d'évaluation fondée sur l'observation directe ne portent pas à l'optimisme. La tentative la plus aboutie a été menée par le « National Board for Professional Teaching Standards » qui a proposé une certification fondée sur l'observation répétée de l'enseignant dans sa classe, et sur le suivi d'un certain nombre de pratiques d'enseignement. Cependant, malgré le coût élevé en temps et en argent de cette procédure, qui dure un an et est relativement sélective (avec un taux d'échec de la moitié), les enseignants certifiés ne semblent que très faiblement plus efficaces que les autres (Goldhaber et Anthony, 2007 [78]). Par conséquent, la mise en œuvre d'une procédure d'évaluation à grande échelle à partir de critères objectifs semble pour l'instant hors de portée.

À défaut d'améliorer leur pratique de recrutement, les systèmes éducatifs pourraient essayer de dispenser une formation efficace aux nouveaux enseignants, que ce soit une formation à la pédagogie, ou bien une formation visant à renforcer des connaissances académiques quand celles-ci s'avèrent insuffisantes. La formation des enseignants fait l'objet d'intenses débats, qui ont souvent trait à son contenu. Ainsi, la formation dispensée dans les IUFM en France avant leur suppression a été fréquemment critiquée à cause de sa supposée abstraction, et du manque de qualifications des formateurs, qui n'avaient fréquemment plus eu d'expérience de l'enseignement dans le primaire ou le secondaire depuis des années.

Ce débat quant au contenu, renvoie à un problème économétrique : il est difficile de mesurer l'impact de la formation des enseignants dans la mesure où il est très difficile de quantifier ce que recouvrent ces formations, qui varient considérablement d'un pays à l'autre. Par conséquent, les leçons tirées d'une expérience nationale sont difficilement transposables dans d'autres contextes.

De manière peu surprenante, les résultats sont d'ailleurs assez divers, mais globalement plutôt négatifs. Ainsi, Angrist et Lavy (2001) [8] obtiennent des résultats positifs de la formation des enseignants sur les résultats académiques pour les écoles primaires laïques, mais pas de résultats pour les écoles primaires privées. L'étude est réalisée à l'aide d'une double différence, où l'on compare l'évolution au fil du temps du groupe d'élèves traité à celle du groupe qui ne l'est pas, ce qui offre de bonnes garanties d'identification. Les formations dispensées sont avant tout des formations à la pédagogie. Cependant, l'étude est menée sur un échantillon relativement restreint. Par ailleurs, le programme de formation semble avoir été associé à d'autres modifications de l'environnement des écoles susceptibles de faire augmenter les scores.

Jacob et Lefgren (2004) [100] étudient un programme de formation des professeurs d'écoles primaires dans la ville de Chicago. Ce programme est spécialement ciblé à destination des écoles comprenant le plus d'élèves en difficulté, et seules les écoles dont les élèves ont des résultats scolaires suffisamment faibles sont éligibles au programme. Le programme se concentre lui aussi sur les pratiques pédagogiques. Jacob et Lefgren comparent les résultats d'élèves venant d'écoles situées juste en dessous et au-dessus du seuil d'éligibilité, donc presque similaires bien que n'ayant pas le même accès au programme de formation. Cette régression sur discontinuité donne, à l'inverse de l'étude de Angrist et Lavy, des résultats nuls du programme de formation des enseignants sur les résultats des élèves. Cependant, le programme de formation étudié portait sur des périodes relativement courtes, ce qui le rend peu susceptible d'être efficace.

De manière similaire, Garet et al. (2008 [73] et 2010) [74] trouvent des résultats nuls dans des études expérimentales où les enseignants sont assignés de manière aléatoire à un programme de formation relativement intensif, aussi bien dans le primaire que dans le secondaire. Harris et Sass (2011) [82] étudient à la fois l'effet de la formation initiale, et celui de la formation continue sur des données américaines (Etat de Floride). Disposant de données longitudinales, ils peuvent contrôler l'effet individuel des étudiants, des professeurs et des écoles par des effets fixes, et estimer l'impact des caractéristiques variant au fil du temps, telles que la formation continue. Harris et Sass trouvent un effet nul ou très faible de la formation continue sur les résultats des élèves, à tous les niveaux d'enseignement et pour toutes les disciplines. Quand des effets positifs existent, ils sont trop faibles pour qu'on puisse considérer socialement rentable d'investir dans la formation des enseignants.

Enfin, il existe également une étude française réalisée par Bressoux, Kramarz et Prost (2009) [33] dont les résultats sont plus contrastés. Bressoux et al. exploitent une erreur de prévision de l'Education nationale française. Sous-estimant le nombre d'enseignants à renouveler dans les écoles primaires, celle-ci n'a pas admis suffisamment d'élèves enseignants dans les écoles normales d'instituteurs qui dispensent une formation de deux ans pour répondre aux besoins. Par conséquent, un groupe d'étudiants supplémentaire a dû être recruté directement sur la liste complémentaire des concours de l'enseignement, sans avoir pu suivre de formation. Cette erreur de prévision a donc créé une

différence aléatoire de durée de formation au sein des enseignants débutants au même moment. Les différences entre ces enseignants sont très faibles, les enseignants recrutés sur liste complémentaire ayant un niveau universitaire quasi similaire à ceux reçus directement.

En comparant les deux groupes d'enseignants, Bressoux et al. trouvent un résultat nul de la formation en français (reading), mais positif et relativement large en mathématiques. Cet effet semble avant tout présent chez les enseignants n'ayant reçu aucune formation en mathématiques durant leurs études universitaires. Par conséquent, dans ce cas précis, il semble que ce soit l'amélioration des connaissances qui a augmenté l'efficacité des enseignants, plutôt que la modification de leurs pratiques pédagogiques. En général, la formation des enseignants semble donc avoir peu d'impact sur la réussite des élèves. Quand cet impact est présent, il semble qu'il puisse être attribué à l'augmentation des connaissances d'enseignants peu formés à certains sujets plutôt qu'à une amélioration de leurs pédagogies. Il semble bien qu'il soit impossible d'enseigner l'enseignement. En revanche, remédier à un manque de maîtrise des connaissances à enseigner peut s'avérer efficace.

## 6.4 Les dispositifs d'incitation des professeurs

A défaut de pouvoir recruter des enseignants de manière efficace, on peut cependant essayer de les rémunérer sur la base de leurs résultats et de leur donner de bonnes incitations. Trois grands arguments peuvent être avancés en faveur d'un salaire conditionné, au moins en partie, à la performance. Tout d'abord, l'efficacité des enseignants peut dépendre d'un effort qui est coûteux pour eux, et pas seulement de caractéristiques fixes. En les rémunérant en fonction de leur efficacité, on les incite à exercer un effort plus important. En outre, rémunérer les enseignants de manière différente peut créer une incitation à quitter l'enseignement pour les enseignants les plus médiocres, et à y rester plutôt que partir dans le secteur privé pour les meilleurs. Enfin, dans la mesure où les individus réagissent à leur perception de la justice, mettre en place un système d'incitation peut constituer une forme de rétribution symbolique pour l'ensemble des enseignants, dès lors que le système d'incitation est perçu comme étant juste. Cela rendrait la profession plus attractive, à l'heure où beaucoup de systèmes éducatifs peinent à recruter suffisamment d'enseignants.

Comme on l'a vu dans la section précédente, un tel dispositif d'incitation est très complexe à mettre en place en raison de la difficulté à disposer de mesures fiables de performance, et surtout de leur forte variance. Une littérature importante en théorie des contrats (Baker et al. , 1994 [19]) montre que l'existence d'un signal ayant une variance importante réduit la possibilité de mettre en place des mesures incitatives puissantes, car elle fait exploser le risque supporté par les agents, qui doivent donc recevoir une compensation supplémentaire.

En effet, dès lors que le signal envoyé par le professeur est très bruité, même un enseignant de grande qualité est susceptible de jouer de malchance (par exemple, en se voyant attribuer une classe où les élèves sont beaucoup plus faibles que leur caractéristiques observables pourraient le laisser penser), et d'être pénalisé. Inversement, un enseignant médiocre peut obtenir une récompense sans effort s'il est chanceux. Par conséquent, la rémunération des enseignants devient susceptible de varier fortement d'une année sur l'autre à cause de facteurs qu'ils ne peuvent contrôler. Or, la plupart des individus ont de l'aversion pour le risque. Il faudrait donc augmenter leur rémunération de manière non négligeable pour leur faire accepter un système d'incitation incluant une part importante d'aléa.

Par ailleurs, un dispositif incitatif peut mener à des comportements visant à contourner le système en satisfaisant les critères du dispositif seulement en apparence. L'un des comportements les plus répandus consiste ainsi à « enseigner pour l'examen » (*teaching to the test*) en enseignant uniquement les compétences nécessaires à la réussite du test, voire la réponse aux questions susceptibles de figurer à l'examen, au détriment d'un apprentissage plus large, plus profond, et surtout plus durable (Carrell et al. 2010 [43]). Par ailleurs, un dispositif incitatif peut aussi amener les enseignants à détourner leur attention des élèves n'ayant que peu de marges de progrès, ou peu susceptibles d'atteindre le niveau minimal requis, au profit des élèves les plus susceptibles de s'améliorer et d'aider le professeur à atteindre ses objectifs. Dans certains cas, un système incitatif peut même mener certains enseignants à tricher et à corriger de manière incorrecte les examens de leurs élèves pour augmenter leurs performances apparentes. Plusieurs études réalisées aux Etats-Unis montrent que cette possibilité n'est pas qu'anecdotique (Jacob et Levitt, 2003 [102]). En outre, les tests utilisés pour mesurer la performance des élèves ne prennent en compte qu'une partie seulement des compétences apprises à l'école. En particulier, comme le montre Jackson (2013) [98], les mesures de valeur ajoutée n'incluent pas la capacité d'un professeur à faire augmenter les compétences non cognitives de ces étudiants (telles que la capacité à fournir un effort, à respecter des règles, et ainsi de suite), alors même qu'elles représentent une part non négligeable des bénéfices tirés de l'éducation.

Par conséquent, un système incitatif mal conçu peut distordre l'effort de l'agent dans une seule direction, ce qui peut être préjudiciable quand il s'agit d'un agent « multi-tâches », devant remplir plusieurs fonctions à la fois. On peut ainsi imaginer que les enseignants se concentrent exclusivement sur les compétences académiques de leurs élèves, négligeant des compétences sociales potentiellement importantes.

Malgré ces difficultés, un certain nombre d'expériences ont été menées aussi bien dans les pays en développement que dans les pays développés. En général, les bonus offerts le sont en addition aux salaires normaux, ce qui réduit le problème du risque supporté par les agents, mais augmente le coût de la rémunération totale. Mais la plupart de ces programmes ont été menés sur une période limitée, ou bien de manière non aléatoire, ce qui rend leur évaluation difficile.

Si l'on se concentre seulement sur les études ayant été réalisées de manière expérimentale, avec une allocation aléatoire des écoles aux dispositifs incitatifs, les résultats sont nuls dans les pays développés. Ainsi, l'évaluation d'un système d'incitation mis en place au niveau du primaire et du secondaire dans l'Etat de New York par Fryer (2013) [71] et Goodman et Turner (2013) [80], trouve des résultats nuls, voire négatifs. Ce système d'incitation avait la particularité d'être fondé sur une récompense allouée non pas à des individus, mais à des groupes d'enseignants affectés à une école donnée, en fonction de l'atteinte d'objectifs fixés au préalable par le district. La répartition entre enseignants se faisait ensuite par un comité formé au sein de l'école. En pratique, la plupart des écoles ont choisi de répartir les primes perçues en fonction du rang hiérarchique des enseignants dans l'établissement, plutôt qu'en fonction d'un critère de mérite. Par conséquent, il apparaît que les incitations fournies sont considérablement diluées par la possibilité, pour un professeur, d'adopter un comportement de passager clandestin. De manière cohérente avec cette idée, Goodman et Turner trouvent que plus le groupe d'enseignants est de faible taille, plus le programme de primes a un effet positif. L'exemple de New-York montre donc l'importance de la manière dont le dispositif d'incitation est construit.

Une autre étude expérimentale conduite à Nashville dans le Tennessee (Springer et al., 2010 [155]) aboutit également à des résultats nuls dans le cadre d'un dispositif incitatif focalisé avant tout sur les enseignants eux-mêmes. Par conséquent, même un dispositif incitatif bien conçu ne semble pas faire preuve d'efficacité.

D'autres études fondées sur des résultats non expérimentaux donnent des résultats positifs. Ainsi, Lavy (2002) [109] étudie-t-il un programme d'incitation mis en œuvre sous la forme d'un tournoi où seules les trois meilleures écoles sont récompensées, de manière analogue au dispositif décrit par Lazear et Rosen (1981) [115], en appliquant une forme de régression sur discontinuité comparant des écoles juste au-dessus et en dessous d'un seuil d'éligibilité pour participer aux programmes (seules les écoles les plus médiocres sont admises à participer). Il trouve un effet assez important de ce programme, bien que ce soit également un programme de groupe comme dans l'expérience new-yorkaise. Dans une étude ultérieure, Lavy (2009) [110], évalue un programme donnant des primes à des enseignants en fonction des résultats de leurs élèves à un examen d'entrée dans l'enseignement supérieur. Les écoles sont assignées au programme en fonction des résultats préliminaires à cet examen, toutes les écoles en dessous d'un certain seuil participant au programme. Cependant, ces résultats préliminaires diffèrent en général de manière importante des résultats finaux, l'écart étant aléatoire. Par conséquent, pour les écoles se situant aux alentours du seuil, la participation au programme peut être considérée comme étant due au hasard. Lavy trouve là encore des résultats positifs relativement élevés. Cependant, ces études portent sur un nombre d'établissements assez restreint (quelques dizaines). Par conséquent, elles restent vulnérables à des chocs affectant des établissements de manière différente selon que l'école est traitée ou non. En ce qui concerne les pays en développement, les résultats des dispositifs d'incitation

évalués par expérimentation aléatoire sont en général positifs (Duflo, Hanna et Ryan, 2012 [57], Muralidharan et Sundararaman, 2012 [131]), bien que les effets de ces programmes ne soient pas toujours durables (Glewwe, Ilias et Kremer, 2010 [77]).

Ces résultats positifs dans les pays en développement pourraient s'expliquer par le moindre contrôle exercé par les directeurs d'écoles et les familles sur les professeurs. Ainsi Duflo et al. documentent-ils un taux d'absentéisme très élevé dans les écoles étudiées au Kenya, de l'ordre de 40

Les résultats des dispositifs incitatifs déjà testés apparaissent donc très mitigés, les études les plus robustes ne trouvant aucun impact dans les pays développés. Il est difficile de privilégier une raison précise de cet échec parmi les explications proposées plus haut. On peut cependant penser que les incitations de groupe, plus faciles à mettre en place lorsque les performances individuelles sont difficiles à mesurer, souffrent d'un problème rédhibitoire de passager clandestin. Par ailleurs, la marge de progression des enseignants dans les pays développés est sans doute limitée.

Ces études sont cependant loin de donner des réponses définitives. En effet, la mesure incitative la plus répandue à travers les systèmes éducatifs est avant tout une évaluation subjective de l'enseignant par un supérieur hiérarchique, qui peut ensuite donner un avis influant sur le salaire de l'enseignant. Ainsi, les enseignants français sont contrôlés de manière régulière, bien qu'espacée, par des inspecteurs. Ils sont également évalués par le chef d'établissement. Ces évaluations qualitatives donnent lieu à une notation qui joue un rôle important dans l'avancement (et donc la progression salariale) des enseignants. Il semble que les évaluations qualitatives soient au moins en partie corrélées à l'efficacité des enseignants (du moins telle que mesurée par leurs effets sur les scores, Jacob et Lefgren (2008) [101], Rockoff et Speroni (2011) [141], Rockoff et al. (2012) [142]).

Il est en pratique difficile de mesurer l'impact de cette rémunération au mérite sur le résultat des élèves, et sur le niveau d'effort des enseignants. Cependant, elle pourrait expliquer l'absence d'effet de mesures incitatives additionnelles. En effet, il est possible que les enseignants soient déjà efficacement contrôlés par leur hiérarchie. Cependant, il faut que cette hiérarchie dispose d'un véritable pouvoir.

Or, plusieurs études ont mis en évidence un effet positif d'une plus grande autonomie des écoles sur les résultats. Ainsi, Clark (2009) [49] évalue un programme visant à donner un statut spécial d'autonomie à une fraction des écoles britanniques. Ce changement de statut dépendant d'un vote des parents compare les écoles où ce vote a été gagné avec une faible marge à celles où il a été perdu avec une marge identique. On peut en effet considérer que pour ce groupe d'écoles le changement de statut a été aléatoire. Clark trouve des résultats positifs de l'autonomie, d'une magnitude relativement large. Un des canaux par lesquels cette autonomie accrue pourrait avoir des effets bénéfiques est un plus grand pouvoir hiérarchique de la direction sur les enseignants. Dans l'étude de Clark, les écoles concernées par le statut d'autonomie se voyaient confier la gestion complète des ressources humaines, et notamment la gestion des salaires, ce qui rend ce

canal plausible.

Quoi qu'il en soit, il semble qu'il y ait peu à attendre de la mise en place de mesures formelles d'incitation. Etant donné le coût non négligeable de ces mesures, leurs avantages potentiels semblent très limités. A l'inverse, des travaux supplémentaires seraient nécessaires pour évaluer l'impact des mesures d'incitation fondées sur une évaluation qualitative des enseignants. Une autonomie accrue des écoles, incluant une plus grande capacité à évaluer et faire varier la rémunération de leurs enseignants, pourrait en effet se révéler une piste peu coûteuse d'amélioration de la gestion des enseignants.

## 6.5 L'appariement entre élèves et enseignants

Outre le recrutement et l'incitation des enseignants, une politique des ressources humaines doit aussi les répartir au sein du système éducatif. En effet, l'équité de la répartition des professeurs au sein des établissements est un sujet relativement sensible. En particulier, les parents d'élèves essaient, autant que possible, d'éviter qu'un enseignant débutant ou peu expérimenté soit assigné à leurs enfants. Par ailleurs, la répartition des enseignants est souvent dénoncée comme injuste. De fait, plusieurs travaux sur le système éducatif américain mettent en évidence un tri important entre les élèves et les enseignants à partir de leurs caractéristiques observables. Les études de Clotfelter et al., vues plus haut, montrent qu'en moyenne les élèves issus de milieux défavorisés ou de minorités ethniques se voient assigner des professeurs moins expérimentés et moins qualifiés. La France n'échappe pas à ce phénomène.

Ce tri entre élèves et enseignants est en général vu comme contribuant à l'existence d'inégalité scolaire entre les groupes par le grand public. On peut être cependant assez sceptique quant à cette idée dans la mesure où seul le fait d'avoir un professeur débutant a un impact important sur les résultats scolaires en primaire, et que la situation au lycée n'est pas très différente, puisque la principale différence est l'existence d'un impact positif de la certification. Par ailleurs, Clotfelter et al. montrent que les différences entre caractéristiques des professeurs ne peuvent expliquer que quelques pour cent de l'écart entre groupes d'étudiants.

Des études plus récentes (Harris et Sass, 2013 [82]) montrent néanmoins un impact un peu plus important de l'expérience au-delà des premières années, que Clotfelter et al. Ces différences entre études s'expliquent par le phénomène d'attrition des enseignants les plus faibles au fil du temps, qui mène à surestimer l'impact négatif du fait d'être un jeune professeur. En comparant la progression de l'efficacité au fil du temps pour un même enseignant, il est possible de mettre en évidence une progression faible mais régulière de l'efficacité dans certaines matières et certains niveaux. Cependant, là encore ces effets sont trop faibles pour que le tri entre élève et enseignant puisse entraîner des inégalités importantes de résultats entre groupes d'élèves.

Un phénomène lié au tri est la rotation importante des professeurs dans les écoles les

plus défavorisées. En effet, les professeurs affectés aux écoles ayant les plus mauvais résultats essaient d'en partir au plus vite en raison des mauvaises conditions de travail qui y règnent. Ce phénomène serait nuisible à la réussite des élèves, en déstructurant les équipes pédagogiques (Ronfelt et al., 2013 [144]).

A première vue, il semble qu'un système centralisé comme le système français puisse aisément manipuler la répartition des professeurs entre les écoles. En effet, il peut choisir directement l'affectation de chaque professeur à un établissement, sans que ce dernier puisse réellement refuser puisqu'il n'existe guère d'alternative à l'éducation nationale pour quelqu'un cherchant un poste d'enseignant. Cependant, ce pouvoir ne s'exerce pas sans limite. Les enseignants ont toujours la possibilité de quitter l'enseignement pour le secteur privé. Par conséquent, nommer les professeurs les plus qualifiés dans les écoles les moins attractives est susceptible d'augmenter leur taux d'attrition.

Par ailleurs, le fait de nommer les professeurs les moins expérimentés dans les écoles les plus défavorisées est aussi une forme de paiement différé les incitant à rester ultérieurement au sein de la profession (de même que la hausse des salaires à l'ancienneté). En effet, affecter les jeunes enseignants aux écoles les plus désirées, et les enseignants expérimentés aux écoles les plus difficiles crée un problème d'incitation. Les enseignants pourraient choisir de changer de profession après leur passage dans les écoles les plus attractives, afin d'éviter la partie de leur carrière passée dans des écoles difficiles. Le système actuel a le mérite de rendre impossible ce comportement en différant la rémunération indirecte associée à l'affectation à une « bonne école ».

Dès lors, même dans un système centralisé, il peut être coûteux de ne pas satisfaire les préférences des professeurs les plus qualifiés et les plus expérimentés. En plus de cet argument d'efficacité, on peut penser que les enseignants ont la capacité d'influencer la répartition des professeurs à travers les organisations syndicales. Celles-ci valorisent en général la rémunération, aussi bien monétairement que non monétairement, à l'ancienneté. Dès lors, changer radicalement la répartition des enseignants entre établissements semble difficile, pour des gains très limités.

Une solution serait d'attribuer des primes importantes aux enseignants expérimentés acceptant d'aller dans les écoles les plus difficiles, et d'y rester suffisamment longtemps. Une telle mesure aurait l'avantage de ne pas remettre en cause le principe de rémunération différée qui permet de limiter l'attrition des enseignants au sein du système éducatif. Le Conseil général de la Seine-Saint-Denis, département qui concentre beaucoup d'établissements défavorisés, et qui peine à retenir des enseignants expérimentés, avait envisagé de mettre en place une telle prime.

Cependant, il n'existe pas, à ce jour, d'évaluations d'une telle politique. En outre, celle-ci pourrait s'avérer coûteuse dans la mesure où il serait nécessaire d'attribuer un complément de rémunération à l'ensemble des enseignants expérimentés en fonction dans les établissements ciblés, même ceux qui seraient restés dans ces établissements de toute manière, faute de pouvoir identifier les enseignants qui sont à la marge de changer d'établissement. Ce genre de mesure génère donc d'importants effets d'aubaine. Néan-

moins, dans la mesure où de telles dépenses contribueraient à augmenter l'attractivité globale du métier d'enseignant, on peut également considérer qu'elles bénéficieraient à l'ensemble du système.

## 6.6 conclusion

En introduction, nous avons évoqué le conflit entre une vision activiste, et une autre plus sceptique, de la gestion des enseignants. Les résultats de la recherche empirique en éducation semblent donner de solides arguments à l'approche sceptique. Ainsi, si les différences d'efficacité entre enseignants sont réelles, les mesurer au niveau individuel est très difficile. Sans mesure précise de l'efficacité des enseignants, les rémunérer au mérite, ou encore les recruter sur la base de leurs capacités potentielles, semble impossible, et peut-être même dommageable. La théorie économique montre en effet que des dispositifs d'incitation mal conçus peuvent se révéler désastreux.

De même, il semble qu'il n'y ait guère à attendre de la formation professionnelle des enseignants qui, la plupart du temps, n'a aucun impact mesurable. Les cours de pédagogie en formation initiale ne semblent pas non plus avoir une quelconque efficacité. Quant aux politiques visant à mieux répartir les enseignants au sein du système éducatif, qui sont relativement consensuelles, elles semblent difficiles et coûteuses à mettre en œuvre pour des gains très limités.

Certaines réformes récentes apparaissent également comme malheureuses. En particulier, recruter des enseignants au niveau master plutôt qu'au niveau licence ne semble pas justifié et entraîne des coûts importants pour les aspirants professeurs, tout en réduisant de manière considérable le vivier de candidats potentiels. De même, réintroduire de nouvelles écoles de professorat (ESPE) alors que la formation à la pédagogie est inefficace semble peu judicieux.

Cela doit-il conduire au pessimisme face à toute réforme, et à l'immobilisme ? De fait, beaucoup de mesures fréquemment proposées semblent dépourvues d'efficacité. D'autres, cependant, n'ont pas encore été suffisamment testées pour pouvoir être réellement évaluées.

Quelques pistes paraissent prometteuses. Tout d'abord, l'existence de fortes différences de performance entre enseignants, différences qui semblent durables et liées à leur personnalité plutôt qu'à leurs efforts, indique que de meilleures procédures de recrutement pourraient avoir des bénéfices importants. Si les mesures de performance fondées sur les scores semblent insuffisamment précises, les combiner avec des évaluations qualitatives, notamment l'évaluation par les pairs et les chefs d'établissement, pourrait se révéler informatif. Il serait donc intéressant d'essayer de manière expérimentale un système de titularisation, où les enseignants ne pourraient être confirmés qu'après deux ou trois années d'enseignement, et où une fraction d'entre eux pourrait se voir remercier à l'issue de cette période d'évaluation. Similairement, donner plus d'autonomie au

chef d'établissement en matière de ressources humaines pourrait s'avérer fructueux, et mériterait d'être expérimenté.

La recherche en économie de l'éducation entraîne la fermeture de certaines portes – en particulier l'espoir un peu naïf de mettre en place des procédures mécaniques d'évaluation et de recrutement fondées sur des scores – mais indique aussi des pistes qui mériteraient d'être explorées. Au scepticisme et à l'activisme, il faudrait sans doute préférer l'expérimentation raisonnée.





## Chapitre 7

Does teachers' compensation reward effectiveness

## 7.1 Introduction

It is largely believed that there is a large agency problem in the educational system, particularly as far as teachers are concerned. Teachers are recruited mostly based on academic credentials, which have little to do with their ability to deliver effective teaching. During the course of their career, they are payed and promoted mainly based on seniority. Moreover, they face very little risk of being fired for being unsuitable for their job. Therefore, many think that merit and effectiveness are not properly rewarded, and that as a result, teachers face little or no incentive to exert more efforts and improve their teaching. This is seen increasingly as a problem, as teachers have been recently shown to be a crucial and long lasting factor in the achievement of students (Chetty et al. 2013 [48]).

This has prompted two reactions. First, many have proposed to introduce more market mechanisms into the educational system in order to put more competitive pressure on schools. If parents have more choice, then they could choose schools which are better at managing and incentivizing teachers. Second, many have proposed to implement pay-for-performance scheme, where teachers are paid according to the relative success of their students in standardized tests.

However, research on the topic has found rather mixed evidence. Studies on school competition have found contradicting findings, and the attempt to implement new incentive scheme often proved disappointing. In general, it seems difficult to design an effective incentive system, because it is difficult to observe the efforts exerted by teachers, who have only a partial control on how their students fare in the exam. Moreover, as shown by Rothstein (2012) [149], implementing an incentive system can be very costly, mostly because it creates a very large risk for teachers.

Research on teachers' incentive has thus mostly focused on newly implemented incentive scheme. However, while seniority plays a very important role in the determination of the pay and promotion of teachers, most existing educational systems includes an evaluation of their teaching staff, and at least partly link teacher advancement and teacher assessment by their hierarchy.

However, we now only a little on the efficiency of these existing incentive scheme. Are they effective in promoting and rewarding the best teachers? It may be that educational system are not as bas as we think in term of incentivizing teachers.

The question is not meaningless since we know that school managers often have information about the quality of their staff . Indeed, subjective evaluation of teachers by their hierarchy often have non-negligible predictive power (Rockoff and Speroni 2011 [141]). As a result, they may have the ability to successfully implement a merit based pay and promotion system.

This question has also interesting implications. If schools are already able to reward the best teachers, it reduces the scope for new incentive scheme to be effective. The failure of recent experiences of implementation of incentives could be due to the fact

that teachers are already incentivized. It could also reduce the interest for market based solution, since it shows that centralized systems are able to manage their human resource effectively.

This study tries to enlighten this question by focusing on a very centralized public education system, the French public secondary education system. Teachers are centrally managed by the administration, with a national pay scale and a national promotion system. The recruitment of teachers is also organized centrally, through national exams, and similarly the geographical assignment of teachers to schools follow national and very rigid guidelines, with very little autonomy at the local level. As a result, it is similar to some large American urban school district, such as the one of New York City or Chicago for instance.

Although public and centralized, the French public secondary education system has an individual component in the remuneration and promotion of teachers, which is not solely determined by seniority rules. Each teacher is ranked according to a number of "steps" ("*échelons*"), which determines his or her wage. The passage to a higher step is determined either by seniority rules, or by a choice by the hierarchy, which enables a teacher to be promoted faster. As a results, two teachers having the same observable characteristics, and in particular the same experience, can be paid differently. However, the difference in term of wage between well and poorly considered teachers is quite small.

However, this could be quite closed to an optimal system, since the performance of teachers is difficult to observe, and that implementing a strong incentive system could make teachers face a large risk, which would make needed to compensate them. The aim of this study is thus to see in what extent this incentive system is able to reward the most effective teachers, defined as the one who are the most able to improve the test score of their students.

To achieve this goal, I focus on four dimensions. First, I look at the relative performance of teachers who have been promoted in a different way in the past, trying to see if teachers who were promoted faster are better than those promoted through seniority rules. Second, I build a measure of rank across teachers having the same level of experience, and I see if it correlates with test score. Third, I look at some objective characteristics which predict efficiency, and I try to see if those having these characteristics enjoy better wage. Last, I measure the variance of the teacher effect, and I see how it compares to the variance measure in other less centralized and more market based system.

Unfortunately, it was not possible to have access to longitudinal data. As a result, I could not reconstitute teacher career, and I had to rely on cross-sectional comparison between teachers apparently similar but having different level of compensation.

The main econometric problem to deal with is the endogeneity of promotion and compensation. There is a potential correlation between compensation and the affectation

to a classroom. For instance, the teachers who are the best considered by their hierarchy could be both promoted and assigned to better students. In the opposite, the teachers who are facing the weakest and the most difficult students could be rewarded by a faster promotion pace.

Another potential econometric problem could be differential attrition. Indeed, teachers can exit, either from the educational system as a whole, or from junior high school to senior high school, or to another status (school director, or "*agrégé*", a status which can be obtained after a difficult examination, and which implies depending on another pay scale, and being assigned to better and older students). This could be a problem if for instance we try to compare teachers having different level of experience, and different ranks, since the best rank teachers tend to exit junior high-school in order to teach senior in high-school, mechanically improving the relative rank of remaining teachers having the same experience.

A last problem is linked to measurement. Indeed, during teachers' career, there is more or less variation among them in term of rank. In the beginning, there is only little difference among teachers. Then, due to the existence of discretionary promotion, more inequalities appear. However, at the end of their career, seniority becomes increasingly important and reduces inequalities. As a result, the scale is not stable, and depend on experience. Another problem is that we do not observe the full career of teachers, but only the date when they are hired. However, it is likely than many of them, particularly women, interrupt their career at some point. Due to these reasons, the assessment of teachers by their hierarchy is imperfectly observed, which could bias estimates toward zero. Last, there is different status, leading to different rules and scale for promotion. To solve the endogeneity problem, I use students' fixed effects. The idea is to compare the results of an identical student across subjects. If being assigned a teacher having some specific characteristics in one subject is associated to better (or worse results) in this subject compared to other subjects, then I can conclude that these characteristics has a positive (or negative) impact on achievement.

The identifying hypothesis is that there is no sorting between students and teachers based on the relative ability of students in one subject conditional on their observables. For instance, there no such thing as a classroom gathering students who are strong in math, and weak in French (conditional on their observables), and which is assigned to an effective math teacher and an ineffective French teacher. This hypothesis can be partially checked by looking at the existence of tracking, and sorting between teachers having specific characteristics and tracks. However, of course, sorting on observable cannot be observed.

To avoid the problem of differential attrition, I build measure of relative rank for teachers having the same level of experience, that is I look at how they fare compared to colleague entered in the educational system at the same time. I also look at different subsamples of teachers having more or less experience.

It is difficult to solve the measurement problem. To alleviate it, I focus on the main

group of teachers, that is "certified" teachers who include more than 80% percent of teachers. I also divide the sample between men and women.

As for the measure of the dispersion of teachers' effects, I compute teacher effects by standard methods, running first students' random effects regression, then computing teachers' fixed effects based on residuals and using precision estimators, and finally shrinking these effects in order to take sampling variance into account.

I find rather surprising results. Teachers who are preferentially promoted are better, but only if they benefit from a strong preference compared to seniority. However, there is a small but significant negative effect of relative rank on efficiency. As for teachers' effect variance, it is similar to what is observed in the United States.

These results seem contradictory. It possible that discretionary promotion favor the best teachers, but that there is a regression to the mean. That is the teachers benefiting the most from "strong" discretionary promotion could then be less favored in term of "weak" discretionary promotion. Moreover, other factors could be considered by the hierarchy in promoting teachers, such as geographical preferences, being assigned to less desired schools, proximity to teacher unions or the hierarchy. Last, all teachers do not enter at the same level. Some can enter the profession at a higher rank, due to their success at the examination. This could explain the absence of link between relative rank and efficiency.

We can then conclude that while knowing at least partially the quality of teachers, education manager fail to implement a merit pay system.

This paper contributes to two strands of literature. First it adds to the literature on pay and promotion policy in educational systems. For instance, Rockoff et al. (2013) [142] investigate how school principals take into account information about teachers' quality to manage they staff. Rockoff and Speroni (2011) [141] looks at the predictive content of subjective assessment of teacher by their hierarchy. It also discusses the implementation of a layoff policy in the New York education system.

Second, it contributes to the literature on incentive scheme in public education systems, with for instance papers of Duflo et al. (2011, 2012) [56] [57] and Muralidharan and Sundararaman (2011) [131] which study incentive schemes in developing countries, and papers by Lavy (2002,2009) [109] [110], Martin (2010) [127], Figlio et al. (2007) [66], Fryer (2013) [71], Goodman and Turner (2013) [80], Glewwe et al. [77], and Springer et al. (2011) [155], which study incentive policy in developed countries.

The rest of the paper proceeds as follow. In the second part, I describe the data, as well as the institutional setting and the pay and promotion procedure. In the third part, I present the identification strategy, and show my results. The last part concludes.

## 7.2 Data and descriptive statistics

### 7.2.1 Students

This study uses administrative data gathered by the DEPP in France, the public office in charge of educational statistics. The data cover the entire French public junior high schools for the year 2005-2010. Since about 80% of French students are going to public school, this provides a fairly comprehensive picture of the French schooling system.

I can observe only one outcome, the "*Brevet des collèges*", which is a comprehensive examination taken at the end of junior high school. It includes both a continuous assessment by the teachers during the two last years of junior high school and a final exam on three main subjects taught in French schools, math, French, and history and geography. The examination is anonymous, graded by teachers who are from different schools than the students, and the questions are identical across the country. As a result, it provides an objective way of assessing student achievement which is not vulnerable to manipulation, as it has been shown in some US states. This is important, since we want to investigate the effect of an incentive measure on efficiency.

I match this data from "*Brevet des collèges*" with administrative data on children, recording an indicator of socio-economic status, the familial situation of the child, if he or she is a repeater or has attended special track of courses inside junior high school, and if he or she received a scholarship. We can also have some information on the nationality of the child, even though this is likely to be misreported, since only very little children are recorded as being of foreign origin.

The table below reports some descriptive statistics on French students. One can notice the large frequency of repeating students in the data. Almost one third of French students have been retained in a grade at least once when they reach the end of junior high school.

---

Table 1 : Descriptive statistics for the public French students  
enrolled in French 9th grade - 2005-2010

---

Variables (N=3,003,616)

---

Percent Professional	0.169
Percent Semi-Professional	0.148
Percent Independent workers	0.084
Percent Agricultural	0.019
Percent White collar	0.166
Percent Blue collar	0.277
Percent Retired	0.018
Percent Inactive	0.094
Percent foreigners	0.036
Percent mother is responsible	0.200
Percent no other adult responsible	0.169
Percent female	0.505
Percent repeater	0.331
Average years late	0.382
Percent Scholarship	0.151
Percent RAR	0.044
Percent RRS	0.144
Percent ZUS	0.088
Percent 'Etablissement sensible'	0.021

---

## 7.2.2 Teachers

Eventually, I match these data with a database recording the characteristics of French teachers. I can observe their age, their experience in the job, their gender, their status and an indication of their assessment by their principal and higher ranked education officials.

They are many different status across teachers. However, most of them are "certified" teachers ("*professeurs certifiés*"). Certified teachers are recruited by a national examination, mostly based on academic knowledge. Other categories of teachers include "*professeur agrégé*". This status can be obtained by passing another exam, considered as being much more difficult than certification, and called "*agrégation*". This exam is often passed after several years of experience by certified teachers. Most "*agrégés*" teach in high schools, but some of them are also affected in junior high schools because there is no post suitable for them in high school. Other category includes "*PEGC*", which are teacher formerly teaching in primary schools and who have been affected to junior high schools, vocational teachers, who teach students belonging to the vocational track, and contractual teachers, who are hired temporarily without passing an exam to fulfill some vacation.

Indeed, as evoked before, teachers are payed according to a nationwide grid, which associates a given salary to a specific rank in the hierarchy. Starting at the beginning of the wage hierarchy in the beginning of their career, teachers can gain "steps" ("*échelon*") until a maximum. Moreover, teachers can be promoted to a higher step in three main ways. First, they can be promoted through seniority, second on recommendation ("*Recommended*"), or on Strongly recommended ("*grand Recommended*"). This gives an indication on their assessment by their hierarchy. I have access to the way by which teachers were promoted to their actual rank (that is, to their most recent step). About 20% of them were promoted through Strongly recommended, 30% through recommendation, and the remaining mostly through seniority.

This promotion are decided by their hierarchy, based on an annual assessment. Assessment can be done in two ways. First, the principal of the school where the teacher is affected gives him or her an annual assessment. Second, teachers are periodically examined by inspectors during in-class visits, which lead to a written assessment. These inspections are however quite rare, one every five to ten years. Based on these assessments, teachers are then more or less quickly promoted.

I have also access to the rank reached by a teacher, that is the number of steps he or she has reached. From this I can build a measure of relative rank across teachers. The main problem is that it is difficult to compare steps across teachers. Indeed, due to seniority rules, the number of steps increase for all teachers through time. It increases faster for the teachers who are the best considered by the hierarchy. However this has an effect mostly on the middle of their career. Indeed, after about 30 years of experience, many teachers reach the maximum number of steps. As a result, there is a compression in the

dispersion of teacher ranks, as the rank is bounded above. Similarly, since all teachers start at the same level in the beginning of their career, there is a compression of the distribution of teachers' rank from below.

Another problem is that many teachers exit. I observe only teachers affected to ninth grade. However, many of them move to high school when they accumulate some experience. Indeed, a post in high school is in general considered as more valuable. As a result, after 15 to 20 years of experience, many of the best ranked teachers leave junior high schools, distorting again the distribution. As a result, a rather well ranked experience teacher can be more weakly considered by his or her hierarchy than an average ranked teacher with little experience, because his well considered colleagues left.

I try to built measure of teacher notation which take these problem into account. The most obvious way to proceed is to center and reduce the distribution of steps for each group of teachers defined by the same year of experience and from the same cohort.

Another possibility is to focus on the relative rank of teacher compared to their cohort. To do so, I first rank them according to the step they have reached, and I then normalize the ranking from 0 to 1. As a result, I obtain a cumulative distribution function, which gives the interval rank to which a teacher belong. I then take the median of this interval. For instance, suppose I can divide teachers into two halves based on their relative ranks, I will then attribute a value of 0.25 to the lower half ( $0.5/2$ ) and 0.75 to the higher half ( $0.5 + (1-0.5)/2$ ). In any case I get a uniform distribution. This enables to somehow compare different cohorts of teachers. However, there is clearly an imperfect observation of the assessment of teachers by their hierarchy, which can be considered as a latent variable.

The table below displays some descriptive statistics on French teachers.

VARIABLES	(1) mean math (sd)	(2) mean French (sd)	(3) mean History (sd)
Experience	12.93 (8.647)	12.72 (8.197)	12.44 (7.314)
Gender	0.503 (0.500)	0.802 (0.399)	0.534 (0.499)
Strongly recommended	0.206 (0.404)	0.198 (0.399)	0.216 (0.411)
Recommended	0.305 (0.460)	0.323 (0.468)	0.353 (0.478)
Promoted through seniority	0.356 (0.479)	0.371 (0.483)	0.339 (0.473)
Promoted for other reasons	0.134 (0.340)	0.107 (0.309)	0.0927 (0.290)
Beginner	0.217 (0.412)	0.203 (0.402)	0.201 (0.401)
<i>“Agrégé”</i>	0.0576 (0.233)	0.0715 (0.258)	0.0625 (0.242)
<i>“PEGC”</i>	0.0596 (0.237)	0.0202 (0.141)	0.0179 (0.132)
Vocational teacher	0.0295 (0.169)	0.0300 (0.171)	0.0281 (0.165)
With a dual certification	0.0152 (0.122)	0.0144 (0.119)	0.0148 (0.121)
With short term contract	0.00539 (0.0732)	0.00374 (0.0610)	0.00351 (0.0591)
Observations	111,308	111,844	108,370

The average level of experience does not differ very much across subjects. It is quite low, reflecting the fact that many teachers tend to move to high school when they gain experience. About 20% of teachers have less than 6 years of experience. Similarly, there is not many differences across subjects as far promotion is concerned. Gender is balanced for math and history, while the very large majority of French teachers are women.

As for teacher status, we can see that certified teachers account for a very large majority of the teaching staff.

## 7.3 Identification strategy and results

### 7.3.1 Identification strategy

As discussed in introduction, the main problem I face is the endogeneity of the assessment by the hierarchy. To solve it, I rely on a specification using students' fixed effects, similar to the one used by Clotfelter et al. (2010) [51]. As a result I compare the results in different subjects of a same students having different teachers. The identification comes from the variation in the characteristics of teachers in different subjects for a same student. As a result, I estimate a very simple specification :

$$\tilde{y}_{ijk} = \tilde{X}_{jk}\beta_k + \tilde{u}_{ijk} \quad (7.1)$$

where  $\tilde{y}_{ijk}$  is the achievement of student  $i$  assigned to classroom  $j$  in subject  $k$  minus the mean of student  $i$  in all three subjects,  $\alpha_i$  is the fixed effect for student  $i$ , capturing his or her underlying ability,  $\tilde{X}_{jk}$  are the characteristics of the teacher of subject  $k$  in classroom  $j$ , minus the mean of the characteristics of all teachers assigned to classroom  $j$ , and  $\tilde{u}_{ijk}$  is the student-subject specific error term. The student fixed effect,  $\alpha_i$ , disappears thanks to the mean differentiation.

I include in  $\tilde{X}_{jk}$  the variable describing the status of the teachers, how they are assessed by their hierarchy, as well as their experience. Since experience mostly matter in the beginning of the career, I model it with dummy variables for each level of experience between 0 to 6 years, and add a quadratic term to account for negative returns to high level of experience.

In order for the estimates to be unbiased, I have to do the hypothesis that students are not assigned to teacher based on their relative strength in one subject compared to the other. For instance, relatively strong student in math compared to French and history should not be systematically assigned to math teachers having specific characteristics. In practice this hypothesis cannot be tested. However, it seems unlikely that schools create subject specific tracks. Indeed, the curriculum is similar for all students. As a result, creating subject specific tracks would not allow to propose a more advanced curriculum to the best students for instance. Moreover, there exists a semi-institutionalized method of tracking in France which is based on the choice of "*options*", that is of courses in an additional subject, most of the time a language. These tracks attract in practice all strong students, without any sort of specific subject preference.

### 7.3.2 Results on the link between pay and efficiency

I first look at the correlation between the way to be promoted at the last step reached and student achievement. This gives the correlation between the contemporary assess-

ment of teachers by their hierarchy and their efficiency. I show below the results both with and without students' fixed effects, on all the classrooms for which I can observe teachers of all subjects.

VARIABLES	(1) Math (OLS)	(2) French (OLS)	(3) History (OLS)	(4) Math (FE)	(5) French (FE)	(6) History (FE)
Experience : 0 year	-0.288*** (0.0886)	-0.0871 (0.0704)	0.326*** (0.0718)	-0.700*** (0.0263)	-0.315*** (0.0206)	-0.491*** (0.0250)
Experience : 1 year	-1.086*** (0.0492)	-0.321*** (0.0398)	0.249*** (0.0387)	-0.671*** (0.0152)	-0.170*** (0.0116)	-0.375*** (0.0135)
Experience : 2 years	-1.025*** (0.0426)	-0.342*** (0.0339)	0.329*** (0.0332)	-0.538*** (0.0130)	-0.0535*** (0.0101)	-0.308*** (0.0115)
Experience : 3 years	-0.942*** (0.0378)	-0.315*** (0.0299)	0.0530* (0.0291)	-0.207*** (0.0116)	0.0428*** (0.00890)	-0.0409*** (0.0102)
Experience : 4 years	-0.656*** (0.0345)	-0.393*** (0.0262)	0.144*** (0.0266)	-0.0355*** (0.0104)	0.0469*** (0.00787)	0.0124 (0.00909)
Experience : 5 years	-0.461*** (0.0332)	-0.0790*** (0.0245)	0.218*** (0.0250)	-0.00330 (0.00983)	0.0398*** (0.00758)	0.00473 (0.00855)
Experience	0.124*** (0.00426)	0.172*** (0.00315)	0.105*** (0.00316)	-0.0451*** (0.00126)	-0.00589*** (0.000970)	-0.0369*** (0.00103)
Squared experience	-0.00156*** (0.000108)	-0.00311*** (7.99e-05)	-0.00171*** (8.36e-05)	0.00104*** (3.17e-05)	5.70e-05** (2.44e-05)	0.000772*** (2.62e-05)
"Agrégé"	1.042*** (0.0262)	0.740*** (0.0179)	0.403*** (0.0186)	-0.0165** (0.00715)	0.0519*** (0.00525)	0.0388*** (0.00588)
"PEGC"	-0.970*** (0.0310)	-1.491*** (0.0361)	-0.855*** (0.0370)	0.169*** (0.0102)	0.180*** (0.00939)	-0.0867*** (0.0105)
With short term contract	-1.767*** (0.0821)	-1.849*** (0.0800)	-1.806*** (0.0822)	-1.247*** (0.0278)	-0.440*** (0.0244)	-1.085*** (0.0316)
Vocational teacher	1.423*** (0.0399)	0.528*** (0.0249)	-4.073*** (0.0256)	0.0697 (0.0617)	0.482*** (0.0665)	-0.583*** (0.0892)
With a dual certification	-0.467*** (0.0488)	-0.300*** (0.0386)	-0.130*** (0.0367)	-0.442*** (0.0138)	-0.0480*** (0.0111)	-0.250*** (0.0117)
Strongly recommended	0.594*** (0.0183)	0.0668*** (0.0133)	0.141*** (0.0131)	0.448*** (0.00527)	0.239*** (0.00402)	0.293*** (0.00437)
Recommended	0.304*** (0.0170)	-0.144*** (0.0122)	0.111*** (0.0120)	0.112*** (0.00485)	0.0539*** (0.00365)	0.00319 (0.00393)
Promoted for other reasons	0.495*** (0.0218)	0.277*** (0.0172)	0.0878*** (0.0177)	0.181*** (0.00643)	0.0549*** (0.00505)	0.0241*** (0.00556)
Gender	0.830*** (0.0121)	0.525*** (0.0114)	0.399*** (0.00886)	0.472*** (0.00364)	0.245*** (0.00312)	0.273*** (0.00304)
Constant	16.67*** (0.0354)	18.12*** (0.0275)	19.47*** (0.0267)	-0.892*** (0.00149)	0.0673*** (0.00125)	0.849*** (0.00125)
Observations	2,425,338	2,427,635	2,427,339	2,405,500	2,405,500	2,405,500
R-squared	0.013	0.011	0.012	0.014	0.005	0.009

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

The effect of promotion on strong recommendation ("strongly recommended") is positive and significant for all subjects, and of the same magnitude once I express it in standard deviation of test scores. While not very large, it is however equal to about one third of the standard deviation of teachers' effects, as it is measured it on the data (results are shown in the next section). By contrast, the effect of promotion on recommendation only is close to null. This is not surprising, given that if about 20% of teachers are promoted on strong recommendation, about 30 % are on recommendation. As a result, these teachers belong to the middle of the distribution, where it is difficult for school managers to distinguish the relative quality of teachers. In the opposite, teachers promoted on strong recommendation should belong to the right tail of ability, which is more easily observable. Therefore, these results give support to the idea that school managers and academic inspectors are able to partially promote teachers based on efficiency.

Sorting appears to be non trivial, since results can change quite much between the simple OLS specification, and the one with students' fixed effects. The change is particularly strong for experience, and for some measure of status. For instance, teachers who are "*Agrégé*" seem to be assigned strong students, even after controlling for observables, since the magnitude of the estimates change very much between the OLS and the fixed effects specification. This is not surprising, since it is well known that these teachers are most of the time assigned to junior high schools located in wealthy area, and gathering strong students. The opposite effect is present for teachers being "*PEGC*" (that is entered in the profession with lower credentials), who are in general assigned to junior high schools in less wealthy area.

Overall, once student unobserved heterogeneity is controlled for by fixed effects, the coefficients are of the expected signs, and comparable to what has been found in the previous literature. Experience has an impact only for beginners, which is quite large. Women are significantly better than men. Teachers having more credentials (that is, "agrégé" who passed a difficult academic exam) do not perform better, and similarly, those with less credentials (that is, "PEGC" who only passed the exam for being primary teachers) do not perform worse. However, teachers recruited for filling temporary vacancy have a strong negative impact in math and history, probably because there is a strong turnover, and because they can be affected to a classroom in the middle of a year to replace a colleague who is unexpectedly missing.

I then look at the correlation between relative rank and student achievement. To do so, I restrict the sample to certified teachers, who are paid and promoted on the same scale and thus more directly comparable. I use the two notation variable, which, as discussed before, capture the relative pay of a teachers compared to his or her colleague having the same experience. I also create variable base on each quartile of the distribution of notation. I show below the fixed effect specification for all subject and all measure of ranks.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Math (FE)	Math (FE)	French (FE)	French (FE)	History (FE)	History (FE)
Experience : 0 year	-0.798*** (0.0368)	-0.840*** (0.0367)	-0.435*** (0.0288)	-0.475*** (0.0289)	-0.780*** (0.0357)	-0.822*** (0.0357)
Experience : 1 year	-0.953*** (0.0186)	-0.981*** (0.0187)	-0.257*** (0.0145)	-0.307*** (0.0146)	-0.551*** (0.0164)	-0.577*** (0.0166)
Experience : 2 year	-0.725*** (0.0154)	-0.724*** (0.0155)	-0.164*** (0.0123)	-0.193*** (0.0125)	-0.443*** (0.0135)	-0.440*** (0.0136)
Experience : 3 years	-0.352*** (0.0138)	-0.279*** (0.0138)	-0.0270** (0.0107)	0.000983 (0.0107)	-0.0938*** (0.0118)	-0.0375*** (0.0119)
Experience : 4 years	-0.105*** (0.0121)	-0.147*** (0.0122)	-0.0169* (0.00934)	-0.0359*** (0.00938)	-0.0509*** (0.0105)	-0.0893*** (0.0106)
Experience : 5 years	-0.0298*** (0.0115)	0.0358*** (0.0116)	0.0369*** (0.00907)	0.0713*** (0.00915)	-0.00708 (0.00979)	0.0417*** (0.00994)
Experience	-0.0331*** (0.00150)	-0.0401*** (0.00154)	-0.00198* (0.00117)	-0.00883*** (0.00119)	-0.0442*** (0.00121)	-0.0500*** (0.00123)
Squared experience	0.000633*** (3.82e-05)	0.000900*** (3.92e-05)	-0.000108*** (2.95e-05)	0.000105*** (3.02e-05)	0.000887*** (3.12e-05)	0.00109*** (3.20e-05)
Gender	0.498*** (0.00434)	0.506*** (0.00434)	0.247*** (0.00379)	0.248*** (0.00379)	0.319*** (0.00357)	0.325*** (0.00357)
notation	-0.0632*** (0.00213)		-0.0377*** (0.00164)		-0.0679*** (0.00174)	
2nd quartile of notation		0.360*** (0.0105)		0.221*** (0.00830)		0.116*** (0.00845)
3rd quartile of notation		0.387*** (0.00989)		0.159*** (0.00798)		0.164*** (0.00804)
4th quartile of notation		0.200*** (0.00981)		0.0658*** (0.00784)		0.00389 (0.00797)
Constant	-0.963*** (0.00176)	-0.964*** (0.00176)	0.0378*** (0.00150)	0.0375*** (0.00150)	0.964*** (0.00145)	0.966*** (0.00145)
Observations	1,606,960	1,606,960	1,606,960	1,606,960	1,606,960	1,606,960
R-squared	0.012	0.013	0.004	0.005	0.008	0.008

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

VARIABLES	(1) Math (FE)	(2) Math (FE)	(3) French (FE)	(4) French (FE)	(5) History (FE)	(6) History (FE)
Experience : 0 year	-0.830*** (0.0344)	-0.840*** (0.0344)	-0.371*** (0.0276)	-0.365*** (0.0276)	-0.817*** (0.0345)	-0.807*** (0.0344)
Experience : 1 year	-0.962*** (0.0185)	-0.878*** (0.0187)	-0.261*** (0.0144)	-0.239*** (0.0146)	-0.573*** (0.0164)	-0.483*** (0.0165)
Experience : 2 years	-0.722*** (0.0154)	-0.666*** (0.0155)	-0.155*** (0.0123)	-0.146*** (0.0125)	-0.445*** (0.0135)	-0.393*** (0.0136)
Experience : 3 years	-0.350*** (0.0137)	-0.406*** (0.0138)	-0.0273** (0.0106)	-0.0361*** (0.0107)	-0.100*** (0.0118)	-0.131*** (0.0119)
Experience : 4 years	-0.103*** (0.0121)	-0.124*** (0.0121)	-0.0184** (0.00932)	-0.0152 (0.00935)	-0.0544*** (0.0105)	-0.0617*** (0.0105)
Experience : 5 years	-0.0296*** (0.0115)	-0.0941*** (0.0116)	0.0343*** (0.00906)	0.0165* (0.00914)	-0.0101 (0.00977)	-0.0566*** (0.00984)
Experience	-0.0324*** (0.00148)	-0.0290*** (0.00149)	-0.00229** (0.00116)	-0.000436 (0.00117)	-0.0449*** (0.00119)	-0.0383*** (0.00120)
Squared experience	0.000612*** (3.75e-05)	0.000577*** (3.78e-05)	-9.70e-05*** (2.91e-05)	-0.000143*** (2.94e-05)	0.000909*** (3.07e-05)	0.000758*** (3.10e-05)
Gender	0.495*** (0.00433)	0.494*** (0.00433)	0.245*** (0.00378)	0.245*** (0.00378)	0.316*** (0.00356)	0.320*** (0.00356)
notation (rank)	-0.264*** (0.00813)		-0.163*** (0.00621)		-0.260*** (0.00656)	
2nd quartile of notation (rank)		0.118*** (0.00622)		0.0128*** (0.00488)		0.154*** (0.00501)
3rd quartile of notation (rank)		-0.0978*** (0.00668)		-0.0414*** (0.00509)		0.0262*** (0.00546)
4th quartile of notation (rank)		-0.145*** (0.00662)		-0.139*** (0.00509)		-0.126*** (0.00533)
Constant	-0.963*** (0.00176)	-0.964*** (0.00176)	0.0378*** (0.00149)	0.0381*** (0.00149)	0.964*** (0.00145)	0.964*** (0.00145)
Observations	1,613,845	1,613,845	1,613,845	1,613,845	1,613,845	1,613,845
R-squared	0.012	0.013	0.004	0.004	0.008	0.010

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

In all specification with students' fixed effect, the effect is negative and quite small, while significant. When I estimate the specification with quartile, I find a surprising non-linear relationship across all disciplines and measures of notation, with teachers at the first and fourth quartile being less efficient than teacher in the second and third quartile.

As discussed before, this could be due to either differential attrition, to distribution compression in the beginning and end of career, or to the fact that female teachers interrupt their career which provoke a mis-measurement of experience. To check for these problems, I run the same regression, but restricted to a sample where all teachers belong to the same window of experience. I choose a 5 to 20 years window, which

should be able to solve the two first problems. I also run regression separately for men and women. Results are presented in the appendix, and do not show any substantial difference with the baseline specification.

As a result, I have rather contradictory results. If efficiency seems rewarded in the short run, this is not the case in the long run.

Next, I try to see if having characteristics positively correlated with students' success in rewarded. To do so, I regress notation variables on gender, while controlling for experience, since it is the only characteristic having a consistent positive effect on score that I know for teachers.

VARIABLES	(1) notation (math)	(2) notation (rank) (math)	(3) notation (French)	(4) notation (rank) (French)	(5) notation (history)	(6) notation (rank) (history)
Gender	-0.0725*** (0.00138)	-0.0267*** (0.000358)	-0.206*** (0.00187)	-0.0606*** (0.000458)	-0.00521*** (0.00137)	-0.00919*** (0.000360)
Experience	-0.000157* (8.82e-05)	0.000123*** (2.09e-05)	-0.000787*** (8.72e-05)	-7.49e-05*** (2.10e-05)	-0.000412*** (0.000101)	0.000139*** (2.43e-05)
Constant	0.0452*** (0.00147)	0.515*** (0.000357)	0.181*** (0.00213)	0.552*** (0.000507)	0.00886*** (0.00153)	0.504*** (0.000377)
Observations	2,119,320	2,123,993	2,200,759	2,202,876	2,154,227	2,156,917
R-squared	0.001	0.003	0.006	0.008	0.000	0.000

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

I find a negative and small but most of the time significant effect. This, however, could be due to the fact that women are more likely to temporarily interrupt their career than men, and thus progress less fast. However, it shows again the educational system fails to reward teachers having characteristics positively associated to performance.

## 7.4 Measuring teachers' effects dispersion

Another way to look at the ability of centralized public education system to manage is to look at the distribution of teachers effects. Indeed, it may be that teachers' management and control is done by other means than by wages. Wage policy could be used for other purposes, for instance in order to reward teachers accepting to go to some geographical area.

A system good at managing teachers should thus be able to compress the distribution of teachers' effects, in particular in the left tails, by compelling the worst teachers either to quit, or to improve compare to their colleagues. As a result, by measuring the

dispersion of teachers' effect, and by comparing it to the one of foreign and more competitive education systems, we could learn about the strength of teacher management and supervision.

### 7.4.1 The first stage equation

In this section, I closely follow the method exposed by Kane and Staiger (2008) for measuring teacher effects. I am interested in modeling the determinant of the test score of the students, in order to recover the causal effect for the teacher. To do this, I decompose the test score of a student  $i$  in a subject  $k$  and in the classroom  $t$  into a part due to the causal effect of a vector of observable  $X_i$ , and three distinct random effect, one for the teacher (in this specific classroom),  $\mu_{kjt}$ , one for the classroom  $\theta_{tj}$ , and one for the student,  $\epsilon_{itjk}$ .

$$Y_{kjt_i} = X_i\beta_k + \phi_{kjt_i} \quad (7.2)$$

where  $\phi_{kjt_i} = \mu_{kjt} + \theta_{jt} + \epsilon_{kjt_i}$

In some specification, I replace the control variable by a student fixed effect  $\alpha_i$ , identified thanks to the variation between the test score of the scores of the students, as in the identification strategy used before :

$$Y_{kjt_i} = \alpha_i + \phi_{kjt_i} \quad (7.3)$$

The advantage of this specification is to control very well for students background. However, it leads to a under-estimation of the variance of the teacher effects, because any correlation between the three subject specific teacher effects will be picked up by the fixed effect. Even tough teacher quality is independent across subjects, natural sampling variation can create correlation at the level of the classroom. If teacher quality is not independent across subject, the downward bias to variance is made even worse.

In order to identify the model, I need to do further assumption.

I make the assumption that  $\mu_{kjt}$ ,  $\theta_{tj}$  and  $\epsilon_{itj}$  are random effects following a normal law with mean 0, and variance equal respectively to  $\sigma_\mu$ ,  $\sigma_\theta$  and  $\sigma_\epsilon$  :

$$\mu_{kjt} \sim N(\mu_{kj}, \sigma_{\mu_{kj}}) \quad (7.4)$$

$$\theta_{kjt} \sim N(0, \sigma_{\theta_t}) \quad (7.5)$$

$$\epsilon_{kjt_i} \sim N(0, \sigma_{\epsilon_{ti}}) \quad (7.6)$$

The distribution of the teacher effect for a given teacher has a mean  $\mu_{jk}$ , and a variance  $\sigma_{\mu_{jk}}$ . The teacher effect for a particular classroom taught by this teacher is drawn from this distribution.

The mean of this distribution is drawn itself from another distribution, the distribution of the mean teacher effect across teachers, which follow the law exposed below :

$$\mu_{kj} \sim N(0, \sigma_{\mu}) \quad (7.7)$$

Each student will have a teacher with a specific  $\mu$ , drawn from a distribution whose mean is drawn from the distribution above.

The first aim of this section is to estimate  $\sigma_{\mu}$ , the variance of the teacher effects. To do so, I will first estimate the variance  $\sigma_{\mu_{kj}}$ ,  $\sigma_{\theta_t}$  and  $\sigma_{\epsilon_{ti}}$ . Using this variance, I will compute the  $\mu_{kjt}$  using Bayes rule, and then, I will directly compute  $\sigma_{\mu}$ .

## 7.4.2 Estimating the variance

In order to identify the three variances I have to make assumptions about the relationship between each of these random effects.

If we consider the covariance between two students being assigned to the same teacher and to different classes, we have :

$$cov(Y_{kjt_i}, Y_{kjt'_i}) = var(\mu_{jk}) + cov(\theta_{t_j}, \theta_{t'_j}) + cov(\epsilon_{kit_j}, \epsilon_{kjt'_i}) \quad (7.8)$$

We assume independence between the  $\theta_{t_j}$  and the  $\epsilon_{kit_j}$ . As a result, we have :

$$var(\mu_{jk}) = cov(Y_{kjt_i}, Y_{kjt'_i}) \quad (7.9)$$

Under the assumption :

$$\theta_{t_j} \perp \epsilon_{kit_j} \quad (7.10)$$

$$\theta_{t_j} \perp \theta_{t'_j} \quad (7.11)$$

I can then identify the variance of  $\epsilon_{t_j}$ . If I subtract the mean of the classroom  $t$  residuals, namely  $\bar{\phi}_t$  from the residual of each student  $i$  of the classroom, I will differentiate

out the classroom effect and the teacher effect, which are identical for all the student of the classroom. As a result, I can write

$$Var(\epsilon_{kjt_i}) = var(\phi_{kjt_i} - \bar{\phi}_t) \quad (7.12)$$

We can define the variance of  $\theta_{tj}$  as the remaining of the variance of  $\phi_{kjt_i}$ , again, on the assumption of the independence between  $\theta_{tj}$ ,  $\epsilon_{kjt_i}$ , and  $(\mu_{jk})$ . As a result, we have

$$Var(\theta_{tj}) = var(\phi_{kjt_i}) - Var(\epsilon_{kjt_i}) - var(\mu_{jk}) \quad (7.13)$$

Under the assumption :

$$\theta_{tj} \perp \epsilon_{kjt_i} \perp \mu_{jk} \quad (7.14)$$

### 7.4.3 Computing the mean teacher effect

I first compute a weighted average of the average classroom residual for each teachers. To do so, I weight each observation by its variance, using a James-Stein estimator (cf. Kane-Staiger 2008). This gives an unbiased minimum variance estimator for  $\mu_{jk}$  . Each classroom average is weighted by its relative precision, which is :

$$w_{jt} = \frac{h_{jt}}{\sum_t h_{jt}} \quad (7.15)$$

where

$$h_{jt} = \frac{1}{Var(\bar{v}_{jt}|\mu_j)} = \frac{1}{\hat{\sigma}_\theta^2 + (\hat{\sigma}_\epsilon^2/n_{jt})} \quad (7.16)$$

The weighted average is then :

$$\bar{\phi}_j = \sum_t w_{jt} \bar{\phi}_{jt} \quad (7.17)$$

( $n_{jt}$  being the number of students per classroom). Thus, classrooms with lower variance and larger size count more compared to classroom with small size and large variance.

Then, the weighted average of the classroom average residuals are weighted according to their reliability, using an empirical Bayes estimator.

$$\hat{\mu}_j = \bar{\phi}_j \frac{\hat{\sigma}_\mu^2}{\text{var}(\bar{\phi}_j | \mu_j)} \quad (7.18)$$

where

$$\text{var}(\bar{\phi}_j | \mu_j) = \hat{\sigma}_\mu^2 + \left( \sum_t h_{jt} \right)^{-1} \quad (7.19)$$

This shrinkage estimator can be directly obtained by the application of Bayes rules, with  $\mu_{kj} \sim N(0, \sigma_\mu)$  being the prior distribution.

#### 7.4.4 The variance of teacher effects

I first present the standard error of the teacher effect obtained by using a shrinkage estimator. The first specification only controls for student characteristics in the first stage. The second specification adds school fixed effects, thus estimating within school teacher effects' variance. The last specification uses students' fixed effects.

---

Table : Standard error of teacher effects computed with a shrinkage correction

---

	(1) Baseline	(2) School Fixed Effects	(3) student Fixed Effects
Math	0.184	0.185	0.110
French	0.108	0.109	0.086
History	0.112	0.055	0.132

---

The estimates are in line with what was found previously in the literature, with a marked difference between the variance of the teacher effects in math and history in the one hand, and French in the other hand. In French and math, the estimates with the specification with OLS or school fixed effects are very similar, and much higher than in the specification with student fixed effects. This may indicate that there is some heterogeneity across classrooms and within school which is not well captured by

observables.

In history, the results are markedly different, with the standard error of the teacher effect being higher in the specification with students fixed effects. This could be due to a specific sorting across students and teachers in history, or the fact that the shrinkage is important in the two first specification, and much less so in the specification with fixed effects.

In order to show the importance of the shrinkage correction, I show below the standard error of the raw fixed effects computed without the shrinkage correction. I can see that the estimates are much higher than with the shrinkage correction. Moreover, the difference between French in the one hand, and Math and History in the other, is much weaker.

---

Table : Standard error of the fixed effects obtained in the first stage regression

---

	(1) Baseline	(2) School Fixed Effects	(3) student Fixed Effects
Math	0.321	0.328	0.194
French	0.253	0.261	0.189
History	0.258	0.186	0.221

---

Eventually, the estimated standard errors appear to be quite important, and point to the existence of a quite large heterogeneity across teachers. Even though teacher effects are not very important compared to other determinant of test scores such that the social background of students, or their personal ability, the fact that they are repeated through time make it an important factor.

At this stage one could be worried about the presence of selection on unobservables, which could take the form of the most (or least) able teachers being assigned the weakest students (for instance, as a strategy to reduce inequalities across students in the first case, or as a way to reward better teachers in the second case). This would lead to understate the standard error of the teacher effects.

Due to the limited numbers of observations for most teachers, it is not possible to

run the usual robustness checks to assess the validity of the computed teacher effects. Indeed, to assess the predictive power of the teacher effects, I need to drop at least one observation by teacher, which reduce a lot our sample. Similarly, it is difficult to perform a test based on teachers move across junior high schools due to the limited number of such movement during the sample period, and due to the fact that I can use only teachers with a high enough number of observation to compute the teacher effects.

We can now try to compare the standard error of teachers effects in France to the one in other countries. If I refer to the most recent and convincing study by Chetty et al. (2014), realized on a large public school district in the US, I find very similar estimates to what is observed in the US. As a result, it seems that level of control is about the same in the French public education system as it is in the American one.

## 7.5 Conclusion

The common knowledge according to which public education system does not reward efficiency nor exert an efficient control on teachers seems to be quite true. This study does not find any consistent positive link between teacher compensation and student achievement. However, it seems that school managers do have some information on teachers quality, and that they use it when they take promotion decision. As a result, it seems that other mechanisms are at work which prevent the best teachers to be the best paid. It may be that some teachers enter the teaching career with some wage advantage for diverse reasons (such as results at the certification exams, or willingness to teach in difficult area), and that they keep this advantage through time, even though they are weaker than the average.

These results support the idea that more incentives could be introduced in the compensation scheme of teachers. In the French case, it seems that the hierarchy partially knows who are the best teachers. As a result, it seems feasible to modify the compensation scheme in order to better reward teacher quality.

## 7.6 Appendix

### 7.6.1 Table 4 : specification with no fixed effects

VARIABLES	(1) Math (OLS)	(2) Math (OLS)	(3) French (OLS)	(4) French (OLS)	(5) History (OLS)	(6) History (OLS)
Experience : 0 year	0.309*** (0.106)	0.407*** (0.106)	0.145* (0.0840)	0.312*** (0.0841)	-0.109 (0.104)	-0.152 (0.104)
Experience : 1 year	-1.208*** (0.0541)	-1.139*** (0.0545)	-0.182*** (0.0429)	-0.0464 (0.0434)	0.0493 (0.0427)	0.0540 (0.0429)
Experience : 2 year	-1.085*** (0.0453)	-1.056*** (0.0456)	-0.340*** (0.0359)	-0.266*** (0.0365)	0.262*** (0.0356)	0.280*** (0.0357)
Experience : 3 years	-1.014*** (0.0397)	-1.113*** (0.0401)	-0.204*** (0.0314)	-0.256*** (0.0314)	-0.0615** (0.0308)	-0.0254 (0.0311)
Experience : 4 years	-0.690*** (0.0361)	-0.617*** (0.0364)	-0.335*** (0.0272)	-0.261*** (0.0273)	0.0685** (0.0280)	0.0180 (0.0282)
Experience : 5 years	-0.418*** (0.0347)	-0.502*** (0.0350)	-0.112*** (0.0256)	-0.212*** (0.0257)	0.169*** (0.0264)	0.189*** (0.0267)
Experience	0.218*** (0.00462)	0.238*** (0.00471)	0.189*** (0.00328)	0.212*** (0.00335)	0.116*** (0.00336)	0.113*** (0.00341)
Squared experience	-0.00413*** (0.000119)	-0.00466*** (0.000121)	-0.00337*** (8.28e-05)	-0.00401*** (8.49e-05)	-0.00194*** (8.98e-05)	-0.00184*** (9.13e-05)
Gender	0.833*** (0.0129)	0.830*** (0.0129)	0.506*** (0.0122)	0.514*** (0.0122)	0.386*** (0.00942)	0.391*** (0.00943)
2nd quartile of notation		0.308*** (0.0331)		-0.162*** (0.0237)		-0.248*** (0.0240)
3rd quartile of notation		0.267*** (0.0310)		-0.0593*** (0.0226)		-0.0618*** (0.0231)
4th quartile of notation		0.566*** (0.0308)		0.269*** (0.0221)		-0.204*** (0.0229)
notation	0.139*** (0.00645)		0.0985*** (0.00471)		-0.0212*** (0.00464)	
Constant	16.33*** (0.0379)	15.83*** (0.0494)	17.98*** (0.0288)	17.76*** (0.0368)	19.47*** (0.0284)	19.64*** (0.0367)
Observations	2,119,320	2,119,320	2,200,759	2,200,759	2,154,227	2,154,227
R-squared	0.014	0.014	0.011	0.011	0.004	0.004

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

VARIABLES	(1) Math (OLS)	(2) Math (OLS)	(3) French (OLS)	(4) French (OLS)	(5) History (OLS)	(6) History (OLS)
Experience : 0 year	0.378*** (0.0997)	0.323*** (0.0997)	0.155* (0.0803)	0.159** (0.0804)	0.138 (0.0978)	0.125 (0.0979)
Experience : 1 year	-1.218*** (0.0539)	-1.126*** (0.0545)	-0.176*** (0.0429)	-0.106** (0.0435)	0.0541 (0.0425)	0.133*** (0.0428)
Experience : 2 year	-1.074*** (0.0452)	-0.986*** (0.0456)	-0.334*** (0.0359)	-0.233*** (0.0368)	0.267*** (0.0355)	0.337*** (0.0358)
Experience : 3 years	-1.011*** (0.0396)	-1.074*** (0.0399)	-0.202*** (0.0314)	-0.226*** (0.0315)	-0.0564* (0.0308)	-0.120*** (0.0310)
Experience : 4 years	-0.688*** (0.0361)	-0.757*** (0.0363)	-0.329*** (0.0271)	-0.352*** (0.0272)	0.0725*** (0.0280)	0.0296 (0.0281)
Experience : 5 years	-0.419*** (0.0346)	-0.473*** (0.0350)	-0.110*** (0.0256)	-0.136*** (0.0258)	0.172*** (0.0264)	0.102*** (0.0267)
Experience	0.219*** (0.00458)	0.218*** (0.00460)	0.189*** (0.00326)	0.193*** (0.00331)	0.118*** (0.00334)	0.120*** (0.00336)
Squared experience	-0.00416*** (0.000117)	-0.00401*** (0.000118)	-0.00339*** (8.22e-05)	-0.00343*** (8.31e-05)	-0.00199*** (8.86e-05)	-0.00199*** (8.95e-05)
Gender	0.832*** (0.0129)	0.825*** (0.0129)	0.504*** (0.0122)	0.496*** (0.0122)	0.386*** (0.00941)	0.388*** (0.00942)
2nd quantile of notation (rank)		0.235*** (0.0193)		0.160*** (0.0145)		0.181*** (0.0139)
3rd quantile of notation (rank)		0.0369* (0.0209)		0.0448*** (0.0149)		-0.0494*** (0.0148)
4th quantile of notation (rank)		0.389*** (0.0204)		0.233*** (0.0150)		0.0298** (0.0145)
notation (rank)	0.463*** (0.0247)		0.298*** (0.0180)		-0.138*** (0.0177)	
Constant	16.10*** (0.0398)	16.14*** (0.0413)	17.83*** (0.0304)	17.82*** (0.0322)	19.53*** (0.0297)	19.37*** (0.0308)
Observations	2,123,993	2,123,993	2,202,876	2,202,876	2,156,917	2,156,917
R-squared	0.014	0.014	0.011	0.011	0.004	0.004

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

## 7.6.2 Table 4 : restricted samples with only teachers from 6 to 20 years of experience

VARIABLES	(1) Math (FE)	(2) Math (FE)	(3) French (FE)	(4) French (FE)	(5) History (FE)
Experience	-0.0655*** (0.00574)	-0.0854*** (0.00582)	0.0254*** (0.00461)	0.0110** (0.00464)	-0.0135*** (0.00464)
Squared experience	0.00208*** (0.000228)	0.00283*** (0.000230)	-0.00117*** (0.000184)	-0.000693*** (0.000184)	-0.000256 (0.000182)
Gender	0.484*** (0.00685)	0.492*** (0.00686)	0.250*** (0.00616)	0.247*** (0.00616)	0.356*** (0.00554)
2nd quartile of notation		0.331*** (0.0167)		0.105*** (0.0135)	
3rd quartile of notation		0.295*** (0.0158)		0.0527*** (0.0130)	
4th quartile of notation		0.0904*** (0.0161)		-0.0717*** (0.0132)	
notation	-0.0829*** (0.00338)		-0.0361*** (0.00270)		-0.0622*** (0.00272)
Constant	-0.757*** (0.00279)	-0.759*** (0.00280)	0.0156*** (0.00244)	0.0167*** (0.00244)	0.780*** (0.00224)
Observations	614,465	614,465	614,465	614,465	614,465
R-squared	0.010	0.010	0.003	0.004	0.009

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

VARIABLES	(1) Math (FE)	(2) Math (FE)	(3) French (FE)	(4) French (FE)	(5) History (FE)	(6) History (FE)
Experience	-0.0652*** (0.00574)	-0.0604*** (0.00574)	0.0249*** (0.00461)	0.0286*** (0.00463)	-0.0141*** (0.00464)	-0.000410 (0.00465)
Squared experience	0.00207*** (0.000228)	0.00196*** (0.000228)	-0.00116*** (0.000184)	-0.00127*** (0.000184)	-0.000233 (0.000182)	-0.000639*** (0.000183)
Gender	0.480*** (0.00686)	0.481*** (0.00687)	0.246*** (0.00617)	0.246*** (0.00617)	0.354*** (0.00554)	0.363*** (0.00555)
2nd quartile of notation (rank)		0.0395*** (0.00925)		0.00560 (0.00737)		0.193*** (0.00732)
3rd quartile of notation (rank)		-0.155*** (0.0103)		-0.0695*** (0.00799)		0.0327*** (0.00814)
4th quartile of notation (rank)		-0.229*** (0.0100)		-0.146*** (0.00784)		-0.0974*** (0.00789)
notation (rank)	-0.343*** (0.0126)		-0.189*** (0.00985)		-0.228*** (0.00996)	
Constant	-0.758*** (0.00279)	-0.759*** (0.00280)	0.0164*** (0.00244)	0.0165*** (0.00244)	0.780*** (0.00224)	0.781*** (0.00224)
Observations	614,481	614,481	614,481	614,481	614,481	614,481
R-squared	0.010	0.010	0.004	0.004	0.009	0.011

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1





## Cinquième partie

# La concurrence entre offreurs d'éducation



## Chapitre 8

Le choix scolaire : un moyen  
d'améliorer l'efficacité du système  
éducatif ?

## 8.1 Introduction

On a vu dans l'introduction que la mise en concurrence des écoles grâce au développement du choix scolaire était l'un des aspects clés des politiques de libéralisation scolaire. Il n'est pas inutile de préciser tout d'abord cette notion de concurrence. La concurrence entre firmes est généralement comprise comme la présence simultanée de plusieurs entreprises sur un marché, ces entreprises essayant d'attirer le plus grand nombre possible de consommateurs afin de maximiser leur profit. De manière analogue, il est possible de définir la concurrence entre écoles comme la rivalité entre plusieurs écoles cherchant à attirer le plus d'élèves possible, ou du moins suffisamment pour pouvoir continuer d'exister. On peut envisager la concurrence entre écoles également comme une compétition pour attirer les meilleurs élèves possible, le but de la compétition étant alors de maintenir sa place ou de l'améliorer au sein d'une hiérarchie scolaire fondée sur la réussite académique.

On le voit donc, la notion de concurrence entre écoles peut prendre plusieurs sens, selon qu'il s'agit de maximiser sa taille, de simplement la conserver, ou bien d'améliorer son rang. En pratique, ces objectifs vont dépendre de la forme organisationnelle prise par les écoles : alors que des écoles à but lucratif chercheront plutôt à maximiser leur taille (pour maximiser leur profit en l'absence de possibilité de discrimination par les prix), des écoles associatives chercheront plutôt à améliorer leur rang et leur prestige. On le voit, la concurrence entre écoles nécessite donc une analyse spécifique, et ne peut être simplement expliquée par des modèles traditionnels d'économie industrielle.

Le but de cette partie est par conséquent de passer en revue les différents résultats empiriques et théorique portant sur les effets de la concurrence entre les écoles. Est-elle capable d'améliorer l'efficacité du système scolaire ? A-t-elle des effets sur les inégalités entre élèves, et sur la répartition des élèves au sein du système éducatif ?

Le but de cette partie n'est donc pas de discuter de l'impact de la libre entrée ou d'une augmentation de la variété lié à une augmentation du choix scolaire, mais uniquement de l'impact pro-compétitif et allocatif de ce choix.

L'existence de gain d'efficacité lié à la mise en concurrence est importante d'un point de vue politique dans la mesure où, comme on le verra ci-dessous, il existe un arbitrage à faire entre les gains potentiels liés à une concurrence accrue, et les inégalités générées par celle-ci. En effet, une libéralisation du marché de l'éducation s'accompagne la plupart du temps d'une différenciation verticale accrue entre écoles. Les gains d'efficacité doivent être larges pour pouvoir justifier une telle hausse des inégalités, qui est perçue par beaucoup comme inéquitable.

Cette partie passera donc d'abord en revue les modèles traditionnels de compétition scolaire, et mettra en évidence quelles sont leur prédictions. Elle décrira ensuite les résultats des études empiriques sur le sujet. Enfin, elle présentera de nouveaux modèles

susceptible de mieux expliquer ces résultats, avant de conclure.

## 8.2 Les arguments en faveur d'un effet pro-compétitif du choix scolaire

Les modèles initiaux de concurrence entre les écoles ont été construits à partir d'une analogie étroite entre écoles et entreprises. De même qu'une entreprise, l'école est supposée vouloir maximiser ses revenus et minimiser ses coûts, afin de les redistribuer à ses membres ou à ses actionnaires le surplus ainsi généré. Elle cherche donc à attirer le plus d'étudiant possible en maximisant son efficacité.

Par exemple, Hoxby (2003) [92], une avocate acharnée de l'augmentation du choix scolaire, modélise la compétition entre écoles comme étant similaire à une compétition à la Bertrand, où une entreprise peut capter l'ensemble du marché si elle offre un prix légèrement inférieur à celui de ses concurrents. Hoxby suppose tout d'abord que l'école ne peut fixer son prix (qui est fixé par le gouvernement). Dans ce cas, les écoles peuvent se distinguer uniquement par la qualité. En supposant une compétition à la Bertrand, les écoles doivent donc maximiser leur qualité sous contrainte de coût pour pouvoir rester sur le marché. Si une école peut fixer son prix, sa qualité peut alors varier, mais conditionnellement à un prix donné, elle maximisera toujours sa qualité.

Hoxby étend l'analyse au cas des écoles à but non lucratif en supposant que ces écoles peuvent redistribuer n'importe quel surplus en direction de leurs employés, que cela soit sous forme d'avantages en nature ou bien de primes. Dans ce cas, bien que atténuées, les incitations sont suffisantes pour répliquer le comportement des écoles à but lucratif, et aboutir à un résultat à la Bertrand.

La compétition aurait donc la capacité de faire atteindre une efficacité maximale au système éducatif. Cependant, cela revient à supposer que la concurrence entre les écoles peut être modélisée comme une concurrence à la Bertrand, ce qui est irréaliste. De plus, on peut être sceptique sur la volonté des employés d'une école de maximiser le surplus à partager (plutôt que de réduire leurs efforts, par exemple). Cette modélisation de la concurrence dans le système éducatif apparaît donc naïve.

Des modèles beaucoup plus sophistiqués ont été développés afin de prendre en compte les spécificités du marché éducatif, et en particulier l'existence d'effets de pairs. Le plus représentatif de ces modèles est celui d'Epple et Romano (1998) [60]. Ces deux chercheurs considèrent une situation où les étudiants varient par revenu et capacité, ces deux dimensions étant positivement corrélées. Ces étudiants valorisent la qualité des autres étudiants fréquentant la même école qu'eux, tout en essayant de réduire leurs frais de scolarité. Par conséquent, chaque école doit prendre en compte l'impact d'un étudiant sur sa qualité lors de sa décision d'admission. Cela signifie que les écoles discriminent par les prix, et offre des prix plus faibles aux meilleurs élèves, internalisant

ainsi l'effet qu'ils exercent sur les autres. Les écoles, privées et à but lucratif, sont libres d'entrer sur le marché, mais font face à des coûts fixes. Il existe également un secteur public gratuit et non sélectif.

Epple et Romano et Romano démontre tout d'abord que dans un tel modèle, l'équilibre sur le marché scolaire est stratifié. Les bons élèves tendent à se regrouper dans les meilleurs écoles. En effet, les élèves les plus riches, qui sont aussi en moyenne meilleurs, sont prêts à payer plus cher pour aller dans les meilleurs écoles. Parallèlement à cela, les meilleurs écoles sont prêts à subventionner les étudiants de bon niveau mais à faible revenu, afin d'améliorer leur qualité. Dans les meilleurs écoles, les étudiants les plus riches et les plus faibles subventionnent donc les plus pauvres et les meilleurs. A l'inverse, les écoles les plus faibles concentre les moins bons étudiants qui ne sont aussi pas suffisamment riche pour payer le surcoût demandé à l'entrée des meilleurs écoles en raison de leur niveau.

Epple et Romano montre également qu'à l'équilibre, l'allocation des étudiants aux écoles est efficace : il n'existe pas d'étudiants prêt à payer plus pour entrer dans une meilleure école. La discrimination par les prix permet aux écoles de s'approprier une partie du surplus du consommateur, pour les étudiants qui ne sont pas marginaux, c'est à dire qui ne sont pas susceptible de changer d'école pour peu qu'une autre école leur propose un prix légèrement inférieur. Tous les étudiants ne peuvent être marginaux, dans la mesure où il existe un nombre fini d'école de qualité différente. Cependant, l'existence d'un profit positif a uniquement une dimension redistributive, et n'entraîne pas de perte d'efficacité.

Epple et Romano simule ensuite l'effet de l'introduction de chèques éducation d'un montant variable. Sans surprise, un montant plus élevé augmente la part du secteur privé. Epple et Romano s'intéresse également à l'aspect redistributif du subventionnement de la demande d'école privée. Il montre qu'une part accrue du privée mène à une augmentation des inégalités en terme de réussite scolaire. En effet, l'existence d'écoles privées amène à concentrer les meilleurs élèves dans les mêmes écoles, et à leur donner une meilleure éducation au détriment des élèves les plus faibles. Dès lors que les familles les plus riches ont une propension à payer plus élevée pour une éducation de qualité, il existe également une augmentation des inégalités d'éducation en terme de revenu. En effet, les familles les plus riches sont à la fois capables et prêtes à payer plus pour faire admettre leurs enfants, même faibles, dans des écoles de qualité. Les meilleurs écoles concentrent donc aussi les élèves les plus riches. Par conséquent, les travaux d'Epple et Romano mettent clairement en évidence de quelle manière une libéralisation du choix scolaire et une expansion du privé peuvent accroître les inégalités entre élèves.

D'autres économistes (Nechyba, 2000 [132], Ferreyra, 2007 [64]) se sont attachés à évaluer empiriquement l'impact allocatif et redistributif de la suppression de la sectorisation et de l'introduction de chèque éducation et du libre choix scolaire. Il montre que de telles réformes mènent à une réallocation profonde des habitants au sein des villes, les familles avec enfant pouvant profiter de prix de l'immobilier plus bas dans des

quartiers relativement pauvres, tout en envoyant leurs enfants dans de bonnes écoles en dehors de leur district. Ces réformes induisent également une forte hausse de la stratification scolaire, de manière similaire au modèle d'Epple et Romano, qui profite avant tout aux élèves les plus doués.

La littérature théorique sur le choix scolaire a donc initialement évolué de manière divergente. Les partisans du choix scolaire se sont généralement référés à des modèles très simples de compétition, sans doute peu adaptés à l'analyse des marchés éducatifs, tandis qu'une autre branche de la littérature, issue de l'analyse des biens publics et des finances locales, s'est avant tout concentrée sur l'impact redistributif du choix.

Si la prédiction selon laquelle la dérégulation du marché éducatif crée plus d'inégalité semble bien fondé théoriquement, les arguments théoriques sur l'impact en terme d'efficience reste donc très généraux et peu développés.

### 8.3 Des résultats empiriques mitigés

De nombreuses études ont essayé d'évaluer l'impact d'un accroissement de la concurrence entre écoles sur la réussite scolaire.

Un premier courant dans cette littérature se concentre sur la concurrence au sein des systèmes scolaires traditionnels. En effet, au sein d'un système sectorisé, les écoles entrent en concurrence dans la mesure où les familles peuvent indirectement choisir une école en choisissant un lieu de résidence. Une école sous performante risque donc de perdre des élèves en provoquant le départ des familles avec enfant vers d'autres districts où les écoles sont meilleures.

Ce mécanisme semble en partie validé par la constatation que les prix immobiliers sont affectés par la qualité des écoles. Une manière simple d'évaluer l'impact de cette forme de choix scolaire sur l'efficacité serait alors de comparer des zones où il est plus ou moins facile de changer de lieu de résidence (par exemple dans de grandes villes denses, ou un déplacement de quelques centaines de mètres peut permettre de bénéficier d'une école différente). Il est néanmoins nécessaire d'instrumenter cette possibilité de choix, qui peut être endogène, dans la mesure où par exemple, les meilleures écoles tendent à absorber les plus faibles, réduisant la densité autour des écoles les plus efficaces.

Hoxby (2000) [90] trouve un effet positif de l'existence d'un plus grand choix en comparant les villes divisées en un très grand nombre de districts scolaires, et celles où un petit nombre de districts gère l'ensemble des écoles. Elle instrumente la fragmentation du système scolaire au niveau métropolitain en utilisant le nombre de rivières, supposées créer des barrières naturelles favorisant la division en district. Cependant, Rothstein (2007) [147] remet ces résultats en question, et montre que l'effet est nul.

Dans une autre étude (2006) [151], il construit un test pour l'existence d'un effet concurrentiel du à un plus grand choix en observant que les districts les plus inégalitaires sont

moins susceptibles d'être affectés par ce choix accru. En effet, les familles se décident en fonction de l'efficacité, mais aussi de la qualité des pairs. Par conséquent, dans un système très inégalitaire, les familles les plus riches ne scolarisent pas forcément leurs enfants dans les écoles les plus efficaces, car celles-ci ne rassemblent pas forcément les meilleurs pairs. Cependant, cette possibilité tend à disparaître quand le nombre d'école augmente, et que les inégalités entre élèves diminuent.

Rothstein propose donc de mesurer l'effet d'un terme d'interaction entre effet de pair et choix sur la réussite scolaire. Si les familles cherchent activement à choisir la meilleur école possible, ce terme doit être positif. En effet, dans un système où le choix est large, les écoles avec les meilleurs pairs sont aussi celles qui sont les plus efficaces, alors que cet effet tend à disparaître quand le choix est faible, puisque les groupes de pair vont avoir un effet déterminant dans le choix des familles.

Rothstein montre que ce terme d'interaction est systématiquement nul, quelque soit la spécification, les contrôles utilisés, ou encore la méthode d'estimation employée (IV ou OLS). Par conséquent, la compétition générée par "le choix de Tiebout", c'est à dire la possibilité pour les familles de choisir leurs biens publics locaux en changeant de lieu de résidence semble inexistante.

Un deuxième courant de la littérature s'est concentré sur des systèmes où le choix est plus développé, mais peut varier de manière importante. Dans ce cas, le mécanisme compétitif ne provient plus du changement de communauté au sein d'un système sectorisé, mais directement du choix d'une école parmi un éventail de possibilité.

L'étude la plus significative est celle menée par Gibbons et al. (2008) [76]. Ces trois auteurs s'intéressent à la concurrence au sein des districts scolaires londoniens. Les familles ont la possibilité de choisir au sein de ces districts, mais pas de choisir une école à l'extérieur de ceux-ci. Par ailleurs, en pratique, il est plus ou moins facile d'aller dans une école au sein d'un district donné pour des raisons de distance. Chaque famille a donc une offre d'école plus ou moins riche, et une plus ou moins grande capacité à se tourner vers la concurrence. En particulier, les familles vivant près des limites de district ont relativement peu de choix. Gibbons et al. peuvent donc construire une mesure de concurrence en regardant quelles sont les alternatives pour la population scolaire potentielle de chaque école. Les écoles dont les élèves ont beaucoup d'alternatives sont soumises à une plus forte pression concurrentielle que les autres.

À partir de cette observation Gibbons et al. peuvent construire une mesure de degré de concurrence pour chaque école. Ils peuvent également instrumenter cette mesure en utilisant la proximité d'une école avec les frontières de district, qui réduit le choix de manière plausiblement exogène. Ils mettent en évidence l'absence d'impact d'un plus grand choix sur l'efficacité des écoles.

Card et al. (2010) [38] utilise une méthodologie similaire pour évaluer l'impact de la concurrence sur les écoles de l'Ontario. Ils utilisent l'existence d'un système confessionnel dual pour obtenir une variation exogène de la concurrence entre écoles. En effet, il existe en Ontario un système scolaire catholique, subventionné par l'Etat mais

n'acceptant que des enfants issus de familles catholiques. Par conséquent, les familles catholiques disposent d'un choix supplémentaire par rapport aux autres, et les zones comportant de nombreux catholiques sont plus concurrentielles. Card et al. utilisent l'évolution de la part de catholique dans une zone et la compare à l'évolution des scores dans cette même zone, afin de supprimer le biais provenant d'une inégale répartition des catholiques en Ontario. Ils trouvent un effet positif mais relativement faible de la concurrence, surtout présent dans les zones nouvellement urbanisées et en expansion, où les parts de marché des écoles ne sont pas stabilisées.

L'impact de variations locales de concurrence semble donc au mieux relativement faibles. Cependant, elles sont étudiées dans des systèmes où le choix est important, mais où les écoles restent à but non lucratif, ce qui peut potentiellement limiter l'impact du choix sur la réussite.

Un troisième courant de la littérature a donc essayé d'étudier l'impact de la mise en place d'un système global de chèque éducation. Étant donné que la seule expérience d'une telle mise en œuvre est limitée au Chili, les seules études sur le sujet porte sur ce pays. Hsieh et Urquiola (2006) [94] combinent plusieurs approches pour évaluer la mise en place de chèques éducation à partir des années 1980. Leur principale stratégie consiste à comparer différentes zones où l'entrée de concurrents privés a été plus ou moins forte, et où leurs parts de marché ont augmenté de manière plus ou moins importante. Ils utilisent des variables instrumentales pour instrumenter le gain de part de marché du privé, et trouve qu'une hausse de la part du privé n'est pas associée à une amélioration des résultats scolaires.

Hsieh et Urquiola compare aussi l'évolution des scores chiliens à ceux de pays économiquement proche grâce aux enquêtes internationales menées à partir des années 1980. Ils montrent que le niveau relatif du Chili tend à se dégrader, ce qui jette le doute sur une politique de chèque éducation. Ces résultats peuvent être également rapprochés des résultats plutôt médiocres de la Suède, qui a expérimenté une politique volontariste de choix scolaire.

## **8.4 Les nouvelles théorisations du marché éducatif : stratification et information imparfaite**

Les test empiriques d'un effet pro-compétitif d'une hausse du choix scolaire se révèle donc nuls, ou seulement faiblement positifs. Ce constat a motivé de nouvelles recherches théoriques sur l'effet du choix sur la réussite scolaire et l'efficacité.

Comme on l'a vu plus haut, Rothstein discute ainsi un modèle où les écoles les plus efficaces ne sont pas forcément choisies par les familles les plus riches, en raison du fait qu'elles peuvent être situées en bas de la hiérarchie en terme de qualité des pairs. Dans un tel modèle, les incitations à exercer un effort sont potentiellement faibles,

et la capacité des écoles les plus efficaces à gagner des parts de marché est limitée. Cependant, un tel modèle considère uniquement des écoles dont l'efficacité est fixe, et qui restent sectorisées. Il est donc trop spécifique pour expliquer les résultats décevants des politiques de libéralisation scolaires.

MacLeod et Urquiola (2012) [123] construisent un modèle modélisant à la fois le niveau d'effort des étudiants, leur réussite scolaire, et leur entrée sur le marché du travail. Le modèle s'intéresse principalement à l'effort des étudiants, et montrent que l'effort est d'autant plus élevé que le système scolaire est peu stratifié. En effet, dans ce cas, être dans une école bien placée dans la hiérarchie scolaire ne donne qu'un signal faible sur ces capacités. Les étudiants sont alors prêts à travailler plus pour acquérir des signaux supplémentaires notamment en passant des examens d'entrée dans des universités sélectives. En effet, c'est le seul signal précis de qualité, et de plus, la faible stratification réduit les inégalités de résultats avant de passer l'examen, et accroît donc la compétition.

MacLeod et Urquiola discutent également de l'entrée de firmes privées sur le marché. Cette entrée a un effet négatif sur l'effort des étudiants dans la mesure où elle accroît la stratification, et donc réduit les incitations à travailler. Par ailleurs, cette stratification a aussi un second effet négatif lié à des problèmes d'allocation. En effet, de manière similaire au modèle développé par Rothstein, l'efficacité des écoles est fixe, et n'est pas liée à leurs groupes de pairs. Par conséquent, des écoles très efficaces mais mal fréquentées peuvent être peu désirées en raison de leur faible groupe de pairs. En entrant sur le marché, les écoles privées peuvent mener à la sortie de ces écoles efficaces, en offrant des moyens de se distinguer aux élèves les plus riches et les meilleurs. De manière plus générale, un secteur privé inefficace peut exister grâce à sa possibilité de sélectionner les étudiants, et réduire ainsi les effectifs d'un secteur public efficace.

MacLeod et Urquiola ne discutent donc pas de l'impact du choix et de la concurrence sur le niveau d'effort ou d'efficacité des écoles, mais traitent seulement de l'effet allocatif de la concurrence. Leur principale contribution est de discuter de l'impact incitatif du libre choix sur les élèves eux-mêmes.

Le dernier article de cette thèse essaiera donc de décrire plus directement les effets incitatifs générés par une augmentation du choix scolaire et de la concurrence entre écoles. Le modèle se concentre sur le cas le plus fréquent de choix scolaire : celui provenant d'écoles à but non lucratif, public ou privée, qui n'ont pas la possibilité de demander des frais d'inscriptions, mais qui peuvent sélectionner leurs élèves. Ce modèle correspond bien à des systèmes scolaires publics qui ont développé le libre choix des élèves entre écoles, tels que, par exemple, la Suède.

Les écoles sont modélisées comme entrant en compétition pour attirer les meilleurs élèves, mais pas pour attirer le plus d'élève possible, leur taille étant fixée. Cette hypothèse peut être justifiée par l'existence de coût fixe important, qui implique l'existence d'un nombre optimal d'élève par école dont il est très coûteux de dévier. Les écoles

entrent donc uniquement dans une compétition pour le statut. Leur qualité provient de leur effort et de leur groupe de pairs. L'effort est coûteux. Cela pourrait refléter par exemple l'idée que mieux contrôler les enseignants impliquent un effort coûteux pour le principal et les parents, ou bien nécessitent des conflits internes que les parents d'élèves souhaitent éviter.

Le modèle compare ensuite une situation où les écoles sont sectorisées, et une autre où le choix est libre et les écoles ont la capacité de sélectionner leurs élèves. Dans le premier cas, la présence d'effet de pair a un effet pro-compétitif. En effet, plus les écoles sont différenciées, plus elles entrent en compétition pour les élèves les plus riches, qui sont les seuls à être suffisamment riches pour être à la fois capables et prêts à changer d'école. À l'inverse, la présence d'effets de pair large réduit le niveau d'effort fourni lorsque le choix d'école est libre. En effet, de forts effets de pairs impliquent une grande différenciation entre écoles. Par conséquent, les élèves préfèrent massivement l'école la plus élevée dans la hiérarchie, et un effort accru ne peut modifier le choix que d'un petit nombre de consommateurs.

Le modèle montre donc que l'introduction d'un libre choix d'école a un effet qui varie en fonction de la force des effets de pairs, et de la manière dont les pairs sont valorisés par les parents d'élèves. Des effets de pairs forts limitent fortement l'impact compétitif du choix, et peuvent même le rendre moins compétitif que lorsque les écoles sont sectorisées. Introduire le choix scolaire peut alors réduire le niveau de compétition entre écoles. Un tel modèle permet donc de rationaliser l'impact faible, voire nul, de l'introduction du libre choix au sein d'un système auparavant sectorisé.

## 8.5 Conclusion

Le marché de l'éducation est un marché marqué par la présence d'effet de pairs et d'information imparfaite. Ces deux caractéristiques ont pour effet d'affaiblir fortement la concurrence entre les offreurs d'éducation. En effet, la présence d'effets de pairs introduit une différenciation verticale entre les écoles, et donc donne un fort pouvoir de marché aux écoles les plus prestigieuses. Quant à l'information imparfaite, elle limite la capacité des élèves à choisir les meilleures écoles, et elle leur donne également une incitation à acquérir des signaux en essayant d'être admis dans les institutions les plus prestigieuses.

Par conséquent, il n'est pas forcément surprenant que l'accroissement du choix scolaire et la libéralisation du marché de l'éducation ne donnent pas lieu à des gains d'efficacité, comme le montre la plupart des études empiriques sur le sujet. Par contre, les effets négatifs pour beaucoup que sont la hausse des inégalités scolaires et la ségrégation des élèves par niveau sont bien documentés.

Dès lors, la prudence s'impose face aux avocats d'une libéralisation du marché éducatif. En effet, il existe sans doute des manières plus efficaces d'améliorer la réussite scolaire

et l'efficience du système éducatif, notamment à travers une plus grande autonomie des écoles, et la mise en place d'incitation à une meilleure gestion.





## Chapitre 9

# School competition and endogenous vertical differentiation

## 9.1 Introduction

Since the seventies, stagnant results and increasing cost have drawn attention to the efficiency of educational system in developed countries. International comparisons (such that the PISA program, led by the OECD) have shown that countries similar in term of wealth and of resources invested in education can differ markedly in term of the academic performance of their students. This has led to a lively debate on how to better use existing resources.

One of the most discussed solution has been to increase competition between schools, in order to provide them with more incentives to perform well. This idea originates from Milton Friedman [70], who observed that schools own a local monopoly, as most of educational systems use a catchment area system, i.e. assign student to schools based on the place where they live. As a result, a non performing school cannot be excluded from the market, since it will enjoy a pool of student living in its catchment area, and unable to switch to a competing school.

Three solutions to this state of affair have been proposed. First, following Friedman, one can entitle all families with a voucher, which enable them to afford a private school. As a result, any school can enter into the market, and parents can freely choose in which school they want to send their children, provided that the voucher is high enough to enable them to afford the school, and that the chosen school is ready to admit their children. A second solution is to facilitate entry into the market for education by allowing parents to create new schools, and by allowing them to benefit from public fund when they reach a critical size. This is well known in the US under the name of "charter school". A third solution consist to suppress the catchment area system, and to let families freely apply to any public or private school.

In the last three decades, these solutions have been widely implemented in developed countries. Thus, New Zealand and Chile have implemented a universal voucher system, leading to a large increase in the private school share of the educational market. Similarly, Sweden has made alternative to the public system largely available by subsidizing private school and by letting families choose freely their school. In the US, recent legislations have tried to encourage the creation of charter schools, and many states have increased school choice inside the public system. In Europe as well, several countries have tried to increase the choice set of families by subsidizing a large private sector, or by increasing school choice.

However, empirical evidence display, at best, mixed results. Indeed, empirical attempts to measure the effects of competition on school efficiency gave most of the time small or insignificant estimates, and sometimes even negative estimates of the effect of competition on test scores (see the review of Urquiola and Macleod, 2012 [124]).

Three strategies of identification have been used. First, some researchers (Hoxby 2000 [90], Rothstein 2006 [151]) have tried to exploit the availability of Tiebout choice, that

is, the possibility for a household to choose the district in which they want to live to benefit from better public goods. They used the density of district in a given metropolitan area as a measure of the easiness of the exercise of the choice of location by families, which enables to measure the intensity of competition. However, these studies, often controversial, tend to point to an absence of effect of competition on school efficiency. Another strategy has been to exploit natural variation in the strength of competition. Such studies include the one of Card et al., 2010 [38], which exploits the existence of a dual schooling system where some schools are sectarians, and some are not, and the one of Gibbon et al. (2008) which uses the proximity to district boundary [76]. However, these studies give only small positive estimate, and only for some subjects or some schools.

A third way to assess the effect of competition is to look at comprehensive voucher reforms. This leads again to disappointing results. Thus, the implementation of school vouchers seems to have led to null and maybe negative effects in the case of New Zealand (Fiske and Ladd, 2001 [67]) and of Chile (Hsieh and Urquiola, 2006 [94]).

This leads to a puzzle : why an apparent increase of competition between schools has only small or no effects on efficiency ? A first explanation is that families do not really care, or are badly informed about the efficiency of school (Rothstein [151]). However, it seems unlikely to be a major factor, since most households seem ready to pay a lot in order to have access to better schools (Black, 2000 [26], Fack and Grenet, 2010 [61]). Another possible explanation is that schools have little scope to improve, because the productivity is more or less fixed in education (for instance, the ability to teach could be an intrinsic ability, little innovation could be done for improving efficiency, and so on). This is at odd with large cross countries difference in achievement, which cannot be easily explained by difference in wealth or by other country specific factors.

Some recent papers have looked at the specific organization of school market to explain why competition between schools could be weak, although the market is left unregulated. Thus, Rothstein emphasizes the fact that the quality of a school does not depend only on its staff and on its management, but also on the students who are attending it. As a result, the best schools in term of quality may not be the best schools in term of efficiency. Since families choose a school mainly by considering quality, this leads to a weakening of the incentive to improve. Similarly, McLeod and Urquiola (2012 [123]) develops of school competition where students select schools based on the past achievement of their graduate, in order to gain a valuable signal on the labor market. They show that this can lead to an inefficient outcome where inefficient schools are the most preferred because of the quality of their student body, which enables them to give a strong quality to their graduates.

These two papers shows that a crucial point in assessing competition between schools is to describe the structure of the market. Indeed, a hidden assumption behind all the empirical studies considered above is that all school are identical, and thus perfectly

substitutable. However, this is obviously not the case, since schools can differ very much with respect to quality. Having to choose between a very bad and a very good school is no choice at all. Therefore, sharp vertical differentiation could lead to a strong market power, and little incentive to improve.

That schools may differ is confirmed by a literature showing that the schooling market tends to be organized in a very peculiar way. Indeed, under the hypothesis that the quality of a school depends at least partially from its student body, and that there is a positive correlation between wealth and student ability, one can show that schools are stratified at the equilibrium, namely that there exists a strict hierarchy of school quality corresponding to a strict hierarchy of social composition of schools. As a result, it shows that vertical differentiation can spontaneously emerge as an equilibrium (Epple and Romano, 1998 [60], Nechyba, 2000 [132], Ferreyra, 2007 [64]).

The deregulation of the school market and the suppression of the catchment area system could lead to a change in the distribution of students among schools, with increasing inequalities. This is a common argument advanced by opponents to school choice programs, who fear that free choice will lead to “cream skimming”, namely to a flight of good students from mediocre public schools to the best one, leading to sharply increased inequalities. From an industrial organization point of view, this would be equivalent to an increase of the vertical differentiation across schools. Provided that the effect is strong enough, it could lead to a weakening of competition.

Last, schools have been generally assumed to only maximize the number of enrolled students, in order to maximize revenue. However, this is unlikely to happen in reality, given that in a stratified market, it would lead to corner solutions where the top ranked schools would be able to attract most students and absorb their competitor, which is not what is observed in reality. Moreover, most schools are not profit maximizers, but are run by a council of parents who care much more about the quality of the students they attract than their quantity, provided they can attract a minimum number of students to cover fixed costs. As a result, competition between schools differs sharply from the traditional model of Cournot or Bertrand competition between firms to which they have been compared.

Building on these observations, this paper tries to build a more realistic model able to incorporate the specific characteristics of the market for school. I explicitly model schools as trying to improve the quality of their student body, rather than trying to increase size at all cost. As a result, I assume that school size is fixed and exogenous. This could be justified, for instance, if parents are reluctant to increase the size of the school to attract new students, while lowering the quality and prestige of the school. The quality of the school from the point of view of the parents is determined both by the average academic quality of its student body, and by the effort of the teachers. This effort is costly for the school. This reflects the fact that teachers are likely to oppose to any attempt to increase their level of effort, which makes it costly for the principal and the parents. Last, rather than modeling schools as being in perfect competition

and being utility takers (Epple and Romano, Epple, Romano and Sieg, 2006 [59]), I assume that schools are strategic, and look for a Nash equilibrium.

I then examine several ways to allocate students to schools. First, I examine a traditional catchment area model, where students are assigned to schools based on where they live. I compare this case with a decentralized allocation process where students can apply to any school, and where schools are free to choose a cut-off level below which they refuse to admit any other student.

I find that the presence of peer effects has a contrasting impact on different type of school system. Surprisingly, large peer effects increase competitive pressure and effort in catchment area systems. Indeed, by increasing stratification, they increase the responsiveness of the school students' quality to effort, since only the richest and ablest student are ready to switch school to enjoy more quality.

Conversely, large peer effects decrease competitive pressure and effort in a school free choice system, where more stratification give schools more market power, and give students less incentive to switch schools to respond to a change in school effort. Under some value of the parameters, these two effects can be strong enough to make the level of effort in a free choice school system lower than in a catchment area school system. As a result, the paper shows how large peer effects can reduce the pro-competitive effect of school free choice. Thus, it proposes a theoretical explanation to why education reforms aiming at introducing more choice and more competition across schools failed to display significant improvement in student achievement. I also propose a simple identification strategy based directly on the model to test its main predictions.

Beyond its contribution to the economics of education, the paper is also a contribution to the study of markets where the consumer is also an input. Such market have been studied first by Rotschild and White (1995) [146], who study the case with perfect competition. Board (2009) [28] study the case with a monopoly. This paper considers an intermediate case where there is imperfect competition. It also investigates the impact of having a non standard objective function for the firm, which corresponds to a competition for prestige or status rather than profit maximization.

The paper proceeds as follow. The second part discusses the previous models of school competition, as well as the features which make them unrealistic. The third part presents the model. The fourth part presents numerical results. The fifth part discusses how the model could lead to empirical predictions, and proposes possible empirical strategies. The last part concludes.

## 9.2 Previous models of school competition

Two different types of model have been developed so far in the literature on school competition. Some empirical studies developed small models, derived from basic industrial organization model. For instance, Hoxby (2003) [92] develops a model where teachers and principals are the residual claimant of any gain induced by an increase in efficiency. Card et al. (2009) develop a model where school compete for attracting a larger number of students.

These models are unrealistic for two reasons. First, they completely ignore the impact of peer effects on the vertical differentiation between schools. As a result, they consider only symmetric games, where school cannot derive market power from having a stronger peer groups.

Moreover, the analogy drawn between schools and firm is dubious. It is unlikely that teachers or principals are school residual claimants. Indeed, most schools are non-profit, and there are very few for-profit actors in the education market. Most schools are directed by a council of parent who would more likely use any additional resources gained from increased efficiency toward further improving the educational quality of the school rather toward monetary redistribution.

Another category of models, much more sophisticated, has been developed specifically to take into account the specificity of the education market. Inspired by the literature on club goods (Rothschild and White, 1995), this literature (Epple and Romano, 1998, Epple, Romano and Sieg, 2006) explicitly models peer effects, and show how they can lead to a stratification of a free education market, with the emergence of an endogenous hierarchy across schools. However, these models make two strong assumptions that are unlikely to hold in practice. First, they assume that schools can perfectly price discriminate across students, in order to take the externality of each student on the others into account. Second, they assume perfect competition : schools are utility takers, and invest all their additional revenues into improving quality. Unsurprisingly, when such model are estimated (for instance, on a sample of private colleges in the paper of Epple et al.), the author finds little market power. They also do not have much to say about the potential efficiency gain from implementing school free choice. Last, they make the implicit assumption that schools (or universities) do not face any agency problem, which is unrealistic.

Another related literature deals with the geographic sorting consequence of different student assignment systems, in particular with the consequences of school vouchers (Nechyba, 2000 [132], Ferreyra, 2007 [64]). These models develop multi-communities models where households have to choose both a house and a school for their children. Then, they investigate the impact of free choice on the allocation of student to schools and on geographical segregation. While giving many insights on the mechanism of segregation and urban stratification, these models are silent on the pro-competitive effect

of school choice.

Inspired by this literature, a few articles have tried to show how the mechanism it describes could prevent parents from choosing the most efficient schools. Thus, Rothstein (2007) shows how, in a catchment area system, the school attracting the best students and the most affluent families may not be the most efficient. Similarly, McLeod and Urquiola (2012) show how households trying to gain a signal of quality for the labor market can prefer inefficient schools because of their peer groups. However, both these papers treat school quality and efficiency as a fixed factor, impossible to modify by the school when facing a changing competitive environment. As a result, they have little to say on the effect on existing schools of an increase in household choice set.

In this article, I try to model realistically school competition in a context where schooling is free, schools cannot charge any additional fees to parents and where schools are non-profit. This would correspond to most public systems in the world, but also to many school systems where private suppliers of education are allowed to operate, but where they are strictly regulated by the state. Sweden and England, which are often seen as examples of countries having liberalized their school market, would fit this description.

A model willing to discuss school competition and the pro-competitive impact of school choice in this context should be able to embody peer effects in the quality function of schools, and should also be able to propose a more realistic school objective function. This last element has been missing in all the previous models, which make them unsuitable to analyze imperfect school competition.

Rather than trying to explicitly model the decision making process within the school, I choose an ad-hoc objective function satisfying two conditions. First, effort is costly for the school. This is obvious from the point of view of the teachers. However, this can also be true from the point of view of the households and the principal. Indeed, increasing the level of effort in a given school would imply to increase the monitoring effort of the teachers, and maybe starting a costly conflict with the staff. I also assume that schools are valuing peer quality, that is, they try to attract the best students. This would also correspond to the quest of prestige. This is a natural goal which is likely to be embraced at the same times by teachers, households, and the principal. Last, I assume that school size is unimportant. This seems to be a reasonable assumption given that parents do not value size (or if they do, rather negatively), but rather school quality, and that teachers and the principal does not derive a specific benefit from increasing size if they are not a residual claimant of any additional funding coming with an increase in size. Moreover, there are a lot of indivisibilities in the cost for schools, and as a result, it is difficult to flexibly adjust size beyond some level (this would imply for instance to open a new class, or to buy new buildings). As a result, schools will stick to a given capacity, and will not try to attract additional students, but rather to increase their admission standards.

I develop in the next part a model of imperfect school competition embodying these features.

## 9.3 A model of imperfect competition between non-profit schools

### 9.3.1 General model

**Households :** There is a continuum of household  $i \in I$ . Each household  $i$  is assumed to maximize its utility function. This one depends on its consumption (which is its income minus the cost of housing in the neighborhood chosen by the household),  $y_i - p_{j(i)}$ , on the amenities of the chosen neighborhood,  $b_{j(i)}$ , on the quality of the school where it send its children,  $q_{k(i)}$ , and on a stochastic term  $\epsilon_k$ , which follows a distribution with cumulative distribution  $F_k$ . The stochastic term represents the idiosyncratic taste for a specific school. All the idiosyncratic shocks are uncorrelated with each other. As a result, the generic utility function of the household is :

$$u_i = U(y_i - p_{j(i)}, b_{j(i)}, q_{k(i)}, \epsilon_{j(i)}) \quad (9.1)$$

I assume that this function is continuous and increasing concave in its three first arguments. Each household is assumed to have only one child. It is defined by a level of income  $y_i$ , and by the intrinsic ability of the child  $a_i$ . We assume that income and ability are correlated, and that they follow a bivariate distribution  $G(y, a)$ , which is continuous in  $a$ , and whose variance of  $a$  does not depend on  $y$ , that is,  $\sigma_a|y = \sigma_a$ .

I also impose a single crossing condition on the income and school quality element of the utility function. This is a standard assumption in order to guarantee the existence of a stratified residential equilibrium. Formally, the condition is :

$$\frac{\partial}{\partial(y_i - p_{j(i)})} \left( \frac{U_2((y_i - p_{j(i)}, q)}{U_1((y_i - p_{j(i)}, q)} \right) = \frac{U_{12}U_1 - U_{11}U_2}{(U_1)^2} > 0 \quad (9.2)$$

This condition means that the marginal utility of school quality is increasing with income, that is, richer parents are valuing more good schools than poorer one. This implies that school quality is a normal good. This condition is important because it will give an incentive to households to sort across communities.

**School quality :** The school quality, as perceived by the parents, depends on the average quality of the students,  $\frac{1}{n_k} \sum_{i \in k} a_i = \bar{a}_k$ , on the level of effort exerted by the faculty and the principal,  $e_k$ . The quality function of the school is then :

$$q_k = Q(\bar{a}_k, e_k) \quad (9.3)$$

The function is continuous and non convex in its first argument, and continuous and increasing concave in its second argument.

**School staff :** The school faculty and the school principal are assumed to maximize an objective function which depends on the average quality of the student body, and their level of effort. The level of effort is bounded and has to belong to  $[\underline{e}, \bar{e}]$  The objective function is :

$$v_k = V(\bar{a}_k, e_k) \quad (9.4)$$

This function is increasing concave in the first and third argument and decreasing convex in the second one. The objective function reflects the fact that schools mostly care about the quality of their student body.

**City :** I assume that there is no vacant house in the city, namely there are as many houses as there are households. The numbers of households and houses is assumed to be fixed. The city is segmented in two different areas, which all have a different level of amenities  $b_j$ , and a single school. Each house within a district  $j$  costs the same price  $p_j$ . This cost could represent as well a rent as interest payments, or local taxes, or any cost linked to residing in this specific district.

**Attendance rules :** We consider two polar cases.

- Each school accepts students based on a catchment area. In order to attend a given school, a child has to live in its area.
- Each school can freely select its students based on their academic result, in order to maximize the quality of its peer group.

**Timing of action :** While setting their level of effort, the schools have to take the other school action into account. As a result, they enter into a strategic game. I model the game between schools in the catchment area model as being simultaneous. However, I model the game in the free choice equilibrium as being a sequential game, where the school starting with the highest peer group (from the catchment area equilibrium) can choose first, followed by the course with the second highest peer group, and so on. This enables to rule out multiple peer equilibrium, and thus mixed strategies. It is also plausible in a context of liberalization, where the most prestigious school can first set their admission procedures, and the other schools follow.

**Resolution and definition of an equilibrium :** I solve this model in two steps. I first determine what is the choice of school and location of the students given school quality, and I compute peer groups for each school given their characteristics and actions and those of their competitors.

In a second step, I solve the problem of the school and obtain a system of non linear equation made of the first order conditions of the school objective function with respect to effort. By solving this system, I can solve the model.

In the case of the catchment area system, the peer groups in each district school are given by :

$$\bar{a}_k = \int_{y_i} \int_{a_i} [F(y_i - p_{j(k)}, b_{j(k)}, \bar{a}_k, e_k) g(y_i, a_i) a_i] da_i dy_i \quad (9.5)$$

with  $F(y_i - p_{j(k)}, b_{j(k)}, \bar{a}_k, e_k)$  being the probability that an individual  $i$  with wealth  $y_i$  choose a school  $k$  with effort  $e_k$  and peer group  $\bar{a}_k$ , located in district  $j$  with amenity  $b_{j(k)}$ .

There are  $K - 1$  such equation (the  $k^{th}$  equation being redundant).

Moreover, there  $J - 1$  size constraints for the district, and  $J - 1$  prices, with  $J$  being the number of district (and of schools in the case of a catchment area system) :

$$n_j = \int_{y_i} [F(y_i - p_{j(k)}, b_{j(k)}, \bar{a}_k, e_k) g(y_i, a_i)] dy_i \quad (9.6)$$

The  $j^{th}$  price is equal to 0, and corresponds to the least desired district. These two sets of equations define at least one residential equilibrium for any vector of school effort level.

Given a residential equilibrium, there are  $K$  derivatives of the objective functions of the schools with respect to  $e_k$  :

$$\frac{\partial V_k}{\partial e_k} = 0 \quad (9.7)$$

Any equilibrium school strategy has to respect this third set of equation.

As a result, a catchment area equilibrium is a vector of  $K$  level of efforts, and  $J - 1$  prices which respects the conditions aboves. Note that at least one price is equal to 0, which reduce the number of equation.

In a second case, there is free choice, but school are size constrained, and can freely choose their students based on ability. In this case, the peer group equation becomes :

$$\bar{a}_k = \frac{F(\bar{a}_k, e_k) \int_{a_i > \underline{a}_k} [g(y_i, a_i) a_i] da_i}{F(\bar{a}_k, e_k) \int_{a_i > \underline{a}_k} [g(y_i, a_i)] da_i} = \frac{\int_{a_i > \underline{a}_k} [g(y_i, a_i) a_i] da_i}{\int_{a_i > \underline{a}_k} [g(y_i, a_i)] da_i} \quad (9.8)$$

where  $\underline{a}_k$  is the minimum threshold level set by the school.

The function  $F(\bar{a}_k, e_k)$  which determines the probability of choosing a school depends now only on the characteristics of this school, that is the peer group  $\bar{a}_k$  and the level of effort  $e_k$ . Since it does not depend from the ability of the student, it can be simplified out.

The objective functions are identical, and the district size constraint is replaced by a school size constraint :

$$n_k = \int_{a_i > \underline{a}_k} F(\bar{a}_k, e_k) g(y_i, a_i) da_i \quad (9.9)$$

An equilibrium in this setting would be a set of  $K$  peer groups,  $K$  level of effort, and  $K - 1$  admission thresholds. Price can be ignored since they have no longer any effect on students allocation to school.

### 9.3.2 Resolution with two schools, a continuum of ability type, and explicit functional forms

In the numerical simulation of the model, I use functional forms which are in use in most of the estimated or calibrated models on school choice and school stratification in the literature.

The school quality function is defined as :

$$q_k = \bar{a}_k^A e_k^B \epsilon_k \quad (9.10)$$

The utility function is defined as :

$$U_{ijk} = (y_i - p_j)^\alpha b_j^\delta (\bar{a}_k^A e_k^B)^\phi \epsilon_k \quad (9.11)$$

A household prefers a school (or a district) to another if :

$$\begin{aligned} \alpha(\log(y_i - p_j) - \log(y_i - p_{j'})) + \delta(\log(b_j) - \log(b_{j'})) + \phi A(\log(\bar{a}_k) - \log(\bar{a}_{k'})) + \phi B(\log(e_k) - \log(e_{k'})) \\ > \log(\epsilon_{k'}) - \log(\epsilon_k) \\ \Leftrightarrow Z_{ijk} > \tilde{\epsilon} \quad (9.12) \end{aligned}$$

I assume that  $\tilde{\epsilon}$  follows a logistic distribution. As a result, the probability for an individual  $i$  to choose school  $k$  and district  $j$  is  $F(Z_{ijk})$ , where  $F(\cdot)$  denote the logistic cumulative density function.

A natural way to model the joint distribution of income and ability would be to use a bivariate normal. However, solving a model including bivariate normals would require to use bivariate numerical integration, which make some computational problem arise. As a result, I prefer to use a discrete distribution for income, and then use a mixture of logistic functions for modeling ability, with the mean of the ability distribution depending on income.

Therefore, I can rewrite the peer group equation in the model with catchment area as follow :

$$\bar{a}_k = \sum_l F(Z_{ijkl})(\mu_a|y_j = y_l)p(y_j = y_l) \quad (9.13)$$

In the previous equation,  $l$  denote a specific income level. The peer group of school  $k$  is equal to the sum of the share of each income group going to the school times the average academic ability in this income group, weighted by the share of this group in the whole population.

I can also define a district size constraint in a similar fashion :

$$n_j = \sum_l F(Z_{ijkl})p(y_j = y_l) \quad (9.14)$$

In the case of a free choice system, the peer group equation of the leading school becomes :

$$\begin{aligned} \bar{a}_1 &= \frac{\sum_l p(y_j = y_l)F(Z_{ijkl})(1 - F(\underline{a})) \int_{\underline{a}}^{+\infty} [f(a|\mu_a(l))a] da}{\sum_l p(y_j = y_l)F(Z_{ijkl})(1 - F(\underline{a}))} \\ &= \frac{\sum_l \left[ p(y_j = y_l)F(Z_{ijkl})(1 - F(\underline{a})) \left( - \left( - \frac{e^{\mu_a(l)} \underline{a}}{e^{\mu_a(l)} + e^{\underline{a}}} - \log(e^{\mu_a(l)} + e^{\underline{a}}) + \underline{a} \right) \right) \right]}{\sum_l p(y_j = y_l)F(Z_{ijkl})(1 - F(\underline{a}))} \end{aligned} \quad (9.15)$$

The school size constraint is :

$$n_1 = \sum_l p(y_j = y_l)F(Z_{ijkl})(1 - F(\underline{a})) \quad (9.16)$$

The peer group of the competitor school can be immediately inferred from the leading school's one.

The objective function of the school is given by a simple polynomial function :

$$V_k(e) = \alpha \bar{a}_k(e) - \theta e_k^\beta \quad (9.17)$$

where  $\alpha$ ,  $\theta$  and  $\beta$  are constant, and  $\beta$  is higher than one.

Given that the function  $\bar{a}_k(e)$  is difficult to characterize, it is needed to use numerical simulation in order to solve the model and obtain comparative statics results. In particular, it is difficult to know how the function  $\bar{a}_k(e)$  will evolve when switching from the catchment area model to the free choice model.

## 9.4 Numerical results and comparative statics

### 9.4.1 Existence of an equilibrium

First, the existence of an equilibrium in the game played between the two schools has to be investigated. For an equilibrium in pure strategies to exist, the reaction function of each of the schools has to be continuous and concave. Since the level of effort is bounded, this has to be true only on the interval defined by the bound.

The respect of this condition depends crucially on  $\bar{a}_k(e_k)$ . This function will be increasing in  $e_k$ . As a result, provided that  $\beta$  is large enough, the objective function of the school will be concave.

Continuity is more difficult to obtain. Indeed,  $\bar{a}_k(e_k)$  may turn out to be discontinuous due to the existence of multiple equilibria in the assignment of students to schools. Two different ways are used to get rid of the problem arising from multiple equilibria.

In the case of the catchment area model, I can make sure that the residential equilibrium is unique by choosing appropriate values for the bound on effort and for the residential amenities. A unique equilibrium where the most preferred school is located in the district with the larger amenities will emerge if two conditions are respected :

- there is no level of effort from the least preferred school which can induce enough students to switch schools such that the least preferred school becomes the most preferred.
- The difference in amenities between the two districts is large enough such that, provided the best school is located in the district with the lowest amenities, there will always be a profitable deviation from the other school such that enough students are willing to switch schools so that schools are exchanging their place in the school hierarchy.

In the case of the free choice model, it is not possible to rule out the existence of multiple equilibria by using appropriate parameters. Indeed, there will always exist a symmetric equilibrium where schools exert the same level of effort and are non-selective. Other equilibria can emerge where one of the two schools is selective. This is possible if the preference for peer is strong enough, and if the upper bound on effort prevents any action enabling the least preferred school to attract students and to reverse school hierarchy.

However, the identity of the most preferred school is undetermined in this case. Indeed, contrary to the residential equilibrium, the difference in amenities does not play any role

any more. As a result, to obtain a pure strategy equilibrium, I have to model the game as being a Stackelberg game where one of the two schools is a leader. This school will in a first step select an admission threshold (the highest one leading to an equilibrium). In a second step, both school will choose their optimal levels of effort.

### 9.4.2 The impact of peer effects on optimal effort : numerical results

In order to clarify the presentation of the result, it is useful to write the first order condition for the two schools :

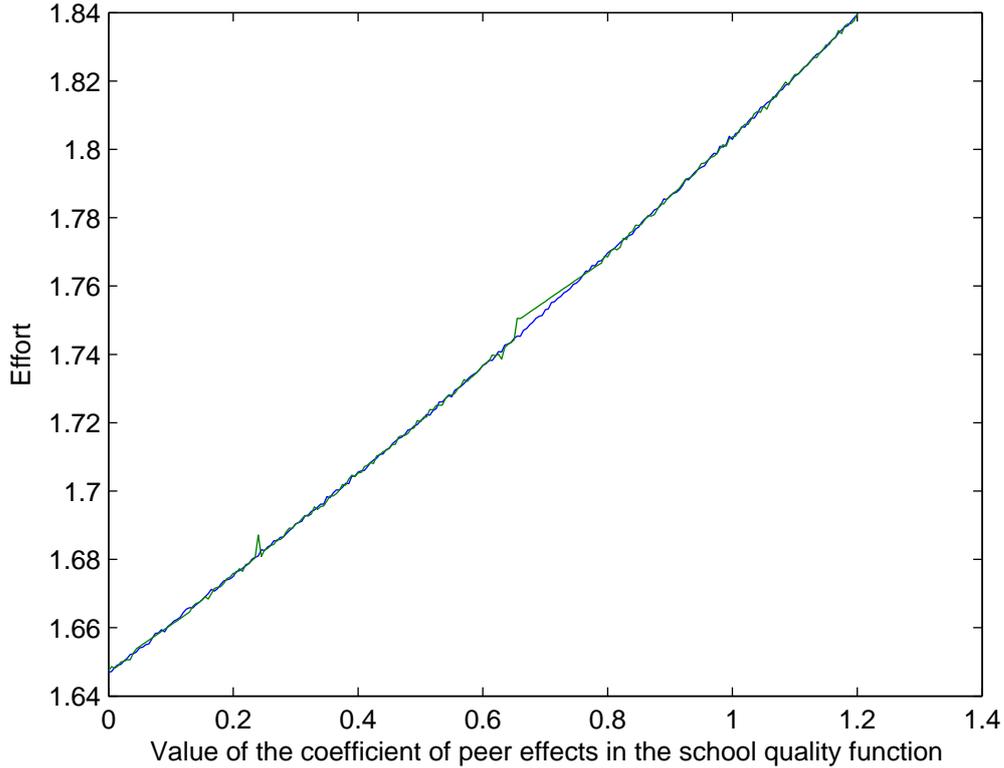
$$\frac{\partial V_k(e)}{\partial e_k} = \alpha \frac{\partial \bar{a}_k(e_k)}{\partial e_k} - \theta \beta e_k^{\beta-1} = 0 \quad (9.18)$$

The level of effort displayed will depend on invariant parameters, that is  $\alpha$ ,  $\theta$  and  $\beta$ . It will also depends on the function  $\frac{\partial \bar{a}_k(e_k)}{\partial e_k}$ , which is the sensitivity of peer groups to effort, and which can widely vary according to market structure and households' preferences. This function will always be positive. The steeper it is, the more incentive schools will have to exert effort.

I first investigate the impact of peer effects on the level of effort displayed by the two schools. When students are assigned to schools through a catchment area system, the impact of school effort on school composition is attenuated by the fact that schools cannot choose their students based on their ability. Instead, households choose jointly a school and a district, according to their level of income and their idiosyncratic shock. Since richer households have a higher propensity to pay for a district with a good school, they are more sensitive to school quality and school effort than poorer households. At the same time, they also have children with a higher average ability. This explains why an increase in effort by a school will increase its peer groups : increasing effort will increase housing demand in the district, and thus make the price increase ; an higher price will increase the average income of the district inhabitants, and thus increase average ability.

The graph below shows the level of effort for a values of  $A$  between 0 and 1.2.

Variation of the level of effort given the importance of peer effects – Catchment area model



The green line represents the value of the effort exerted by the incumbent school, and the blue one the value of the effort exerted by its competitor. While very close, the two lines do not exactly coincide, and the game is not symmetric. Indeed, while the two objective functions are identical, the function  $\frac{\partial \bar{a}_k(e_k)}{\partial e_k}$  need not be constant, which explains why the incumbent schools and its competitor school face different incentives (this would be the case only if the function  $\bar{a}_k(e_k)$  was linear).

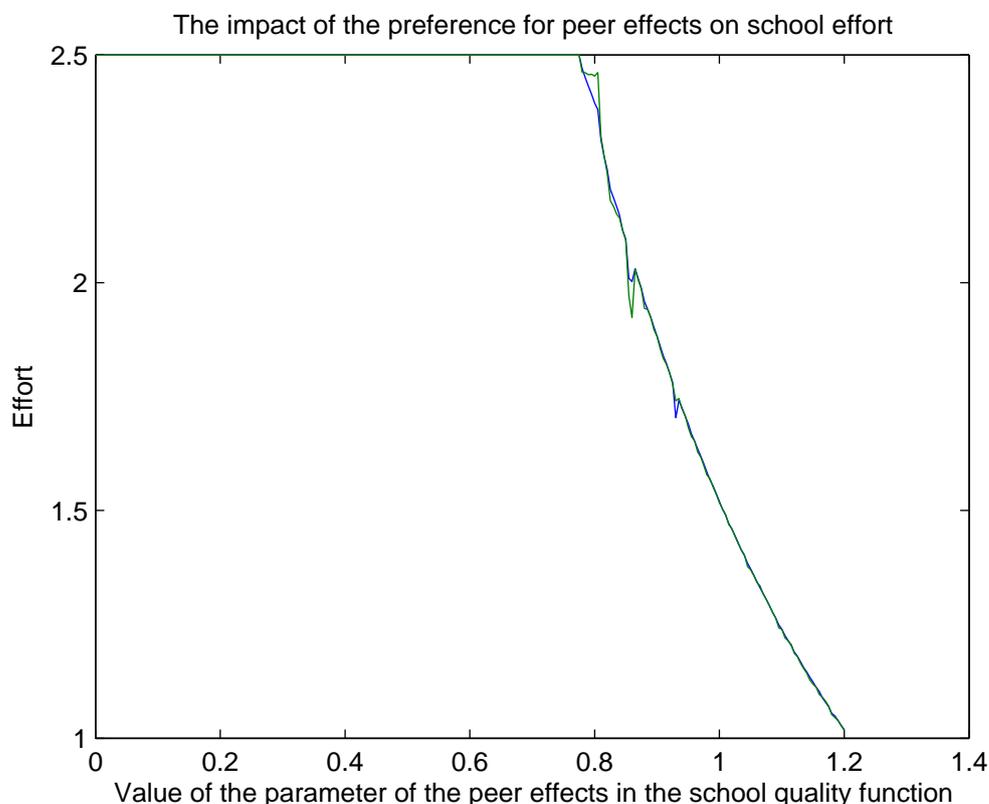
**Numerical result 1 :** In the catchment area model, stronger peer effects increase the level of effort for both schools

This result seems surprising at first. Indeed, one would rather think that by increasing differentiation across schools, peer effects would reduce the amount of competition. The intuition is that in a more stratified school system, the difference in prices across areas are larger, and as a result, the households attending the best school are in average richer. As a result, the household which are on the edge of switching schools when effort increases are richer, and therefore have abler children. Moreover, the presence of peer effects implies the existence of a multiplier effect on effort. Indeed, by increasing the quality of the peer groups, the school additionally further attracts households, and make the price of the district increase even more.

I then look at the impact of peer effects on the level of effort when students are allocated using free choice and full selectivity. In this case, the return on effort is larger, because schools are able to select the best among the students that they attract. As a result, an

increase in attractiveness is rewarded by a larger improvement in the peer group of the school. This explains why the level of effort displayed with free choice is most of the time higher than with a catchment area system (at least for the value I have selected for the simulation).

The graph below shows the level of effort for a values of  $A$  between 0 and 1.2.



The green line represents the value of the effort exerted by the incumbent school, and the blue one the value of the effort exerted by its competitor. Again, the game is not symmetric, and as a result, the level of effort displayed is slightly different for each school. Starting from the upper bound, the level of effort sharply falls with the strength of peer effects. The fall is large enough for the effort in the free choice system to be lower than in the catchment area system, for the value of  $A$  approximatively higher than 1.

**Numerical result 2 :** In the free choice model, stronger peer effects decrease the level of effort for both schools.

This results is directly driven by two factors. First, it is influenced by the shape of the distribution of ability. Since it is a bell-shaped distribution, it implies that attracting more candidates will enable to increase the admission threshold by less and less as the threshold get closer from the mean of the ability distribution. This involves a strong decreasing return to effort as stratification increases. As a result, by increasing stratification and increasing the admission threshold of the best school, stronger peer

effects will reduce effort.

Another important factor is that the idiosyncratic term in the choice equation follow a logistic function. As a result, when the two schools are very differentiated, few students will be willing to switch school due to a change in effort, since only students with a large (and unlikely to happen) idiosyncratic shock will be close of being indifferent between the two schools. As a result, stronger peer effects and more stratification will increase the market power of the leading school by reducing the elasticity of demand with respect to effort.

## 9.5 Discussion and empirical prediction

The model makes various predictions which have empirical application. In a catchment area system, it implies that more stratification will actually increase competition and effort. Conversely, it shows that in a free choice system, more stratification will lead to a decrease in competition and effort.

Taken together, these two predictions point to a differentiated effect of free choice across different type of community. In a community where segregation and stratification are important, because people deeply care about peers, the gain of free school choice are very much reduced. This could explain why many studies on the impact of school choice on achievement fail to find large, or even significant positive effect of competition.

This set of prediction provide us with a simple way to test the model. Testing the model in cross-section could be difficult, since there could be many confounding factor which could correlated with the degree of stratification across areas. However, the model implies that the competitive effect of a reform introducing school choice should widely vary across areas with different level of stratification and segregation. As a result, a possible specification would be to regress the change in academic achievement following the reform on the change on school stratification induced by the reform (measured, for instance, by the change in inequality across schools) :

$$\Delta\bar{a}_{area} = \beta_1\Delta Var(\bar{a}_{school})_{area} + \beta_2\Delta X_{area} + \epsilon_{area} \quad (9.19)$$

$\Delta\bar{a}_{area}$  is the change in average academic achievement in a given area.  $\Delta Var(\bar{a}_{school})_{area}$  is the change in the variance across school of academic achievement in a given area, and  $\Delta X_{area}$  is the change in a set of control variables. If parents in a given area put a lot of weight on school peer quality in their assessment of the overall quality of the school, then  $\Delta Var(\bar{a}_{school})_{area}$  will be large. Indeed, parents will react strongly to free choice by preferring most of the time the schools at the top of the hierarchy, and stratification by ability will strongly increased compared to a catchment area system. If the predictions of the model are valid, the coefficient  $\beta$  should then be negative. Most of the possible

confounding factor should be eliminated by the differentiation, which will remove all the time invariant area specific factors.

However, it is important to understand the limitation of the model. Indeed, the model restricts schools to compete only for better peers. As a result, it does not explore the competitive pressure which could come from competing for the quantity of students in a context where school are on the verge of closing due to being run under capacity. It also ignores the case of for-profit schools, which do not care about quality for itself, but only want to maximize profit, and may want as a result to attract weak students.

Therefore, this models highlights the fact that policies aiming at increasing school competition should not only aim at expanding school choice. They should also change the way school are run, and probably encourage the entry of for-profit suppliers of education. However, this is a step that many would disagree with for moral and political reasons.

## 9.6 Conclusion

This paper developed a model of imperfect competition between two non-profit school competing for student quality. This model is more realistic than those developed earlier in the literature, because it uses a realistic objective function for schools, and that it features a prominent role for peer effects and imperfect competition. Numerical simulations show that the presence of peer effects has a strong impact on the level of effort exerted by schools at the equilibrium. While peer effects have a small positive effect on effort in a catchment area system, they strongly decreased the level of effort in free choice system. As a result, strong peer effects reduce the potential benefit of school free choice very much.

These findings may explain why empirical studies failed most of the time to find a positive effect of more school choice on student achievement. It lead also to empirical prediction and to a natural way of testing them.

## 9.7 Appendix

### 9.7.1 Values used in the numerical simulation

I use the following (arbitrary) values in solving the model :

- Parameters of the joint distribution of income and ability, and parameters of the districts' amenities
- $\mu_y=10$  ; (Average of the income distribution)

- $\mu_{y_{min}}=4$ ;
- $\mu_{y_{sup}}=16$ ;
- $\sigma_y=1.5$ ;
- $\mu_x=7$ ; (Average of the ability distribution)
- $\mu_{x_{min}}=3$ ;
- $\mu_{x_{sup}}=11$ ;
- $\sigma_x=0.5$ ;
- $\rho=0.5$ ;
- $b_1=2$ ;
- $b_2=1$ ;
- Parameters of the utility function of the households
  - $\alpha=0.4$ ;
  - $\delta=0.2$ ;
  - $\phi=0.4$ ;
  - $\sigma=0.05$ ;
- Parameters of the quality and objective functions of the schools
  - $A=0.8$ ;
  - $B=0.2$ ;
  - $\beta=0.5$ ;
  - $\xi=2$ ;

## 9.7.2 Method of resolution of the model

### Method used to numerically solve the catchment area model

- Algorithm for solving the residential model for any value of  $e$ 
  - compute the value of  $\bar{a}$  given  $e$  and  $p$  by using Gauss-Newton method
  - use bisection method to solve for equilibrium  $p$  given  $e$ .
- Use grid search for finding equilibrium given that the two response function are concave in  $e$

### Method used to numerically solve the free choice model

- Algorithm for solving the residential model for any value of  $e$ 
  - compute the value of  $\bar{a}$  given  $e$  and  $a_{min}$  (closed form)
  - use bisection method to solve for optimal  $a_{min}$
- Use grid search for finding equilibrium given that the two response function are concave in  $e$

The main problem in the two cases is that the peer group  $\bar{a}$  is only approximated with an error. As a result, any attempt to compute a derivative based on it may lead to very imprecise results. In particular, it is not possible to use Gauss-Newton or the steepest

descent method, which rely on derivatives, to solve for the equilibrium effort of the two schools. This is why I need to use grid search to find the optimal level of effort.





## Sixième partie

### Conclusion

# 1 Une tentative de bilan

L'éducation est un sujet sensible et central, tant il touche à l'avenir d'une société et à l'image qu'elle veut se faire d'elle-même à travers ses enfants. Les débats sur les questions éducatives sont donc inévitablement vifs, polémiques, et souvent, idéologiques. Fort d'une expérience personnelle, en tant qu'élève, parent ou enseignant, beaucoup ne se soucie pas d'établir des faits, et défendent des convictions qu'ils considèrent comme fondamentales. La question du rôle de l'Etat et du marché dans le système éducatif est encore plus idéologique car elle croise à la question éducative celle de la place de l'Etat, qui reste un facteur fort de division idéologique.

Le but de cette thèse aura donc été d'essayer d'apporter des faits et des analyses rigoureuses dans un débat où le ton est souvent péremptoire. A lecture de la littérature économique, et suite aux travaux originaux présentés dans cette thèse, que peut-on conclure sur l'impact de l'introduction des mécanismes de marché dans le système éducatif?

La première partie de cette thèse a discuté la manière dont le marché pouvait mener à un meilleur appariement entre élèves et écoles. Elle aboutit à la conclusion que la présence d'effets de pairs, d'imperfection de l'information, et de rationalité limitée rendent peu probable une capacité supérieure du marché à allouer les étudiants aux écoles. En particulier, le chapitre deux montrent que l'utilisation d'algorithme d'appariement adapté peu permettre d'améliorer le bien être social en réduisant la stratification entre école et en réduisant la corrélation entre préférences. Le chapitre trois montrent quant à lui que les décisions d'admissions d'une grande école peuvent être inefficaces, et que certains étudiants entreprennent des études alors même que leur chance de réussite sont faibles. La situation émergeant spontanément d'une interaction sans entrave entre écoles et élèves a donc peu de chance d'être la meilleure.

La seconde partie de cette thèse a mesuré la présence d'effets de pairs au niveau du collège en France. Ces effets de pairs apparaissent comme plutôt larges, et générés avant tout par la présence de redoublants, c'est à dire à travers la présence des élèves les plus faibles, ce qui est cohérent avec la littérature existante. Cela confirme donc la présence d'effets de pairs et d'interactions entre élèves sur le marché de l'éducation.

La troisième partie de cette thèse s'intéresse à la manière dont un système public est capable de rémunérer les enseignants en fonction de leur capacité à faire progresser les élèves. Ils montrent des résultats en apparence contradictoires. Si les enseignants promus de manière préférentielle sont en moyenne meilleurs, il n'y a pas de lien sur le long terme entre salaire et capacité. Le système éducatif étudié n'est donc pas capable de mettre en œuvre une rémunération au mérite. Par conséquent, il existe bien un potentiel d'amélioration au sein des systèmes publics. Cependant, cela n'implique pas que les écoles privées soient plus capables de mieux rémunérer les meilleurs enseignants.

Enfin, la quatrième partie discute de l'impact d'un marché scolaire dérégulé sur la productivité et le niveau d'effort des écoles. Comme le montre le chapitre 8, les études empiriques existantes sont globalement décevantes et peine à trouver un impact d'un plus grand choix, ou d'une apparente plus grande pression concurrentielle sur la réussite scolaire. Motivé par ce fait, le chapitre 9 développe un modèle de concurrence entre écoles pour la qualité des élèves et montre que la présence d'effets de pair réduit fortement les gains à attendre en terme de pression compétitive d'une libéralisation du marché scolaire. Sous certaines valeurs des paramètres, une telle libéralisation peut même mener à une baisse du niveau d'effort, en raison de la hausse de stratification entre écoles qui renforce le pouvoir de marché des écoles les plus prestigieuses. Le modèle propose donc une explication à l'absence d'effet positif d'un accroissement du choix scolaire.

La thèse amène donc à réfuter la plupart des arguments mis en avant par les partisans de l'introduction de mécanismes de marché au sein du système éducatif. Un système public centralisé est ainsi plus efficace pour apparier élèves et écoles. Par ailleurs, la compétition entre école sur un marché dérégulé est faible, en raison de la différenciation verticale entre école. La thèse ne permet par contre pas de trancher sur la supériorité en terme de gestion des écoles privées. Elle montre par contre que les écoles publiques sont peu capables de payer leurs enseignants au mérite. Néanmoins, la littérature sur les écoles privés dans les pays développés ne semblent pas montrer une plus grande efficacité de celles-ci.

Par conséquent, on peut conclure, en réponse à la question posée dans l'introduction, que l'introduction de mécanismes de marché dans le système éducatif ne permet pas d'améliorer les résultats scolaires, ni l'efficacité des dépenses d'éducation. Par contre, cela est susceptible de mener à une hausse des inégalités entre étudiants, et à une plus grande ségrégation au sein du système scolaire. Le marché n'est donc pas une solution aux problèmes éducatifs.

## **2 Implication pour les politiques publiques et pistes de recherches futures**

Les avocats d'une libéralisation du marché de l'éducation apportent donc une réponse peu satisfaisante à une question pertinente : comment s'assurer que les fonds investis par l'Etat et les familles dans l'éducation soit utilisés efficacement ? La thèse amène donc à déplacer le débat. Plutôt que des mécanismes de marché, ce sont plutôt de meilleurs mécanismes de contrôle qui sont nécessaires. En effet, les systèmes publics semblent peu capable de résoudre les problèmes d'agence existant au sein de toute organisation de grande taille, comme le montre la troisième partie de la thèse.

Comme le discute cette partie, il n'est pas aisé de mettre en place des systèmes d'in-

citation auprès des enseignants. Cependant, un certain nombre de pratiques semblent susceptible d'améliorer la situation. Tout d'abord, il semble crucial de mettre l'accent sur les procédures de recrutement des enseignants. Ceux-ci ont une efficacité qui varie fortement, et qui semble peu évoluer au cours du temps. Par conséquent, réussir à recruter de meilleurs enseignants est susceptible d'améliorer les résultats scolaires sur le long terme. Dès lors, on peut envisager de mettre en place un processus de titularisation plus progressif, réalisé sur deux ou trois ans, et qui permettrait de rassembler un plus grand nombre d'information sur les aspirants enseignants. Du point de vue de la recherche, il serait intéressant d'essayer de trouver le moyen le plus efficace pour résoudre ce problème d'information auquel les écoles doivent faire face.

Un autre aspect de la gestion des écoles susceptibles d'être amélioré est leur autonomie, et plus généralement leur structure organisationnelle. En effet, les écoles doivent concilier les intérêts de plusieurs parties prenantes (élèves, familles, enseignants, direction, hiérarchie) qui ne sont pas forcément cohérents entre eux. Cela génère un système d'incitation complexe qui est mal compris. L'une des difficultés à faire une analogie entre école et entreprise est ainsi l'impossibilité de définir clairement qu'elle est l'objectif d'une école, alors que l'on peut de manière raisonnable faire l'hypothèse que les entreprises cherchent à maximiser leur profit.

Dès lors, une piste de recherche intéressante, qui serait aussi un moyen d'améliorer le fonctionnement du système éducatif, serait de comprendre comment les décisions sont prises dans les écoles, comment les problèmes d'agence y sont résolus, et comment les objectifs y sont déterminés. Une telle compréhension permettrait de concevoir des mécanismes d'incitation plus à même d'assurer que les écoles remplissent les fonctions qui leurs sont assignées par la société.

Le système éducatif, qu'il soit organisé sous la forme d'un marché ou bien d'une organisation intégrée, a donc des spécificités qui en font un objet d'étude particulièrement intéressant pour les économistes. En effet, d'un point de vue théorique, sa compréhension nécessite de développer de nouveaux outils et concepts dans le domaine de la conception de mécanismes (*mechanism design*) et de la théorie des contrats. D'un point de vue empirique, il offre de larges bases de données permettant de réaliser aisément des études économétriques. L'économie de l'éducation ne manque donc pas de sujet de recherche intéressant ayant des possibilités d'application concrète.





# Bibliographie

- [1] Daniel Aaronson, Lisa Barrow, and William Sander. Teachers and student achievement in the chicago public high schools. *Journal of Labor Economics*, 25(1) :95–135, 2007.
- [2] Atila Abdulkadiroglu. School choice. *Oxford Handbook of Market Design*. Oxford University Press, New York, pages 138–169, 2013.
- [3] Atila Abdulkadiroğlu, Nikhil Agarwal, and Parag A Pathak. The welfare effects of congestion in uncoordinated assignment : Evidence from the nyc high-school match. *Unpublished manuscript*, 2014.
- [4] Atila Abdulkadiroğlu, Parag A Pathak, and Alvin E Roth. Strategy-proofness versus efficiency in matching with indifferences : Redesigning the nyc high school match. *The American Economic Review*, pages 1954–1978, 2009.
- [5] Atila Abdulkadiroglu and Tayfun Sönmez. School choice : A mechanism design approach. *The American Economic Review*, 93(3) :729–747, 2003.
- [6] Armen A Alchian and Harold Demsetz. Production, information costs, and economic organization. *The American economic review*, pages 777–795, 1972.
- [7] Andreas Ammermueller and Jörn-Steffen Pischke. Peer effects in european primary schools : Evidence from the progress in international reading literacy study. *Journal of Labor Economics*, 27(3) :315–348, 2009.
- [8] Joshua D Angrist and Victor Lavy. Does teacher training affect pupil learning? evidence from matched comparisons in jerusalem public schools. *Journal of Labor Economics*, 19(2), 2001.
- [9] Kate L Antonovics and Richard H Sander. Affirmative action bans and the “chilling effect”. *American law and economics review*, pages 252–99, 2013.
- [10] Peter Arcidiacono, Esteban Aucejo, Patrick Coate, and V Joseph Hotz. Affirmative action and university fit : evidence from proposition 209. Technical report, National Bureau of Economic Research, 2012.
- [11] Peter Arcidiacono, Esteban Aucejo, Andrew Hussey, and Kenneth Spenner. Racial segregation patterns in selective universities. *Journal of Law and Economics*, 56(4) :1039–1060, 2013.

- [12] Peter Arcidiacono, Esteban M Aucejo, Hanming Fang, and Kenneth I Spenner. Does affirmative action lead to mismatch ? a new test and evidence. *Quantitative Economics*, 2(3) :303–333, 2011.
- [13] Peter Arcidiacono, Esteban M Aucejo, and V Joseph Hotz. University differences in the graduation of minorities in stem fields : Evidence from california. Technical report, National Bureau of Economic Research, 2013.
- [14] Peter Arcidiacono, Esteban M Aucejo, and Ken Spenner. What happens after enrollment ? an analysis of the time path of racial differences in gpa and major choice. *IZA Journal of Labor Economics*, 1(1) :1–24, 2012.
- [15] Peter Arcidiacono, Gigi Foster, Natalie Goodpaster, and Josh Kinsler. Estimating spillovers using panel data, with an application to the classroom. *Quantitative Economics*, 3(3) :421–470, 2012.
- [16] Peter Arcidiacono and Michael Lovenheim. Affirmative action and the quality-fit tradeoff. Technical report, Duke Working Paper, 2014.
- [17] Ian Ayres and Richard Brooks. Does affirmative action reduce the number of black lawyers ? *Stanford Law Review*, pages 1807–1854, 2005.
- [18] Ben Backes. Do affirmative action bans lower minority college enrollment and attainment ? evidence from statewide bans. *Journal of Human Resources*, 47(2) :435–455, 2012.
- [19] George Baker, Robert Gibbons, and Kevin J Murphy. Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics*, 109(4) :1125–1156, 1994.
- [20] Coralio Ballester, Antoni Calvó-Armengol, and Yves Zenou. Who’s who in networks. wanted : the key player. *Econometrica*, 74(5) :1403–1417, 2006.
- [21] Patrick Bayer, Randi Hjalmarsson, and David Pozen. Building criminal capital behind bars : Peer effects in juvenile corrections. *The Quarterly Journal of Economics*, 124(1) :105–147, 2009.
- [22] Gary S Becker. *Human capital : A theoretical and empirical analysis, with special reference to education*. University of Chicago Press, 2009.
- [23] Eric P Bettinger, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu. The role of application assistance and information in college decisions : Results from the h&r block fafsa experiment\*. *The Quarterly Journal of Economics*, 127(3) :1205–1242, 2012.
- [24] Stephen B Billings, David J Deming, and Jonah Rockoff. School segregation, educational attainment, and crime : Evidence from the end of busing in charlotte-mecklenburg. *The Quarterly journal of economics*, 129(1) :435–476, 2014.
- [25] Anders Bjorklund, Melissa A Clark, Per-Anders Edin, Peter Fredricksson, and Alan B Krueger. *The market comes to education in Sweden : an evaluation of Sweden’s surprising school reforms*. Russell Sage Foundation, 2005.

- [26] Sandra E Black. Do better schools matter? parental valuation of elementary education. *Quarterly journal of economics*, pages 577–599, 1999.
- [27] Nick Bloom and John van Reenen. Measuring and explaining management practices across firms and countries. *Quarterly Journal of Economics forthcoming*, 2007.
- [28] Simon Board. Monopolistic group design with peer effects. *Theoretical Economics*, 4(1) :89–125, 2009.
- [29] Vincent Boucher, Yann Bramoullé, Habiba Djebbari, and Bernard Fortin. Do peers affect student achievement? evidence from canada using group size variation. *Journal of applied econometrics*, 29(1) :91–109, 2014.
- [30] John Bound, Michael Lovenheim, and Sarah Turner. Why have college completion rates declined? an analysis of changing student preparation and collegiate resources. Technical report, National Bureau of Economic Research, 2009.
- [31] Yann Bramoullé, Habiba Djebbari, and Bernard Fortin. Identification of peer effects through social networks. *Journal of econometrics*, 150(1) :41–55, 2009.
- [32] Yann Bramoullé, Rachel Kranton, and Martin D’Amours. Strategic interaction and networks. *The American Economic Review*, 104(3) :898–930, 2014.
- [33] Pascal Bressoux, Francis Kramarz, and Corinne Prost. Teachers’ training, class size and students’ outcomes : Learning from administrative forecasting mistakes. *The Economic Journal*, 119(536) :540–561, 2009.
- [34] William A Brock and Steven N Durlauf. Discrete choice with social interactions. *The Review of Economic Studies*, 68(2) :235–260, 2001.
- [35] Thibault Brodaty. Les effets de pairs dans l’éducation : une revue de littérature. *Revue d’économie politique*, 120(5) :739–757, 2011.
- [36] Mary A Burke and Tim R Sass. Classroom peer effects and student achievement. *Journal of Labor Economics*, 31(1) :51–82, 2013.
- [37] Antoni Calvó-Armengol, Eleonora Patacchini, and Yves Zenou. Peer effects and social networks in education. *The Review of Economic Studies*, 76(4) :1239–1267, 2009.
- [38] David Card, Martin D Dooley, and A Abigail Payne. School competition and efficiency with publicly funded catholic schools. *American Economic Journal : Applied Economics*, 2(4) :150–176, 2010.
- [39] Katherine Grace Carman and Lei Zhang. Classroom peer effects and academic achievement : Evidence from a chinese middle school. *China Economic Review*, 23(2) :223–237, 2012.
- [40] Scott E Carrell, Richard L Fullerton, and James E West. Does your cohort matter? measuring peer effects in college achievement. *Journal of Labor Economics*, 27(3) :439–464, 2009.

- [41] Scott E Carrell and Mark L Hoekstra. Externalities in the classroom : How children exposed to domestic violence affect everyone’s kids. *American Economic Journal : Applied Economics*, 2(1) :211–228, 2010.
- [42] Scott E Carrell, Bruce I Sacerdote, and James E West. From natural variation to optimal policy? the importance of endogenous peer group formation. *Econometrica*, 81(3) :855–882, 2013.
- [43] Scott E Carrell and James E West. Does professor quality matter? evidence from random assignment of students to professors. *Journal of Political Economy*, 118(3), 2010.
- [44] Stephanie Riegg Cellini, Fernando Ferreira, and Jesse Rothstein. The value of school facility investments : Evidence from a dynamic regression discontinuity design. *The Quarterly Journal of Economics*, 125(1) :215–261, 2010.
- [45] Yeon-Koo Che and Olivier Tercieux. Efficiency and stability in large matching markets, 2013.
- [46] Raj Chetty, John N Friedman, and Jonah E Rockoff. Measuring the impacts of teachers i : Evaluating bias in teacher value-added estimates. Technical report, National Bureau of Economic Research, 2013.
- [47] Raj Chetty, John N Friedman, and Jonah E Rockoff. Measuring the impacts of teachers i : Evaluating bias in teacher value-added estimates. Technical report, National Bureau of Economic Research, 2013.
- [48] Raj Chetty, John N Friedman, and Jonah E Rockoff. Measuring the impacts of teachers ii : Teacher value-added and student outcomes in adulthood. Technical report, National Bureau of Economic Research, 2013.
- [49] Damon Clark. The performance and competitive effects of school autonomy. *Journal of Political Economy*, 117(4) :745–783, 2009.
- [50] Charles T Clotfelter, Helen F Ladd, and Jacob L Vigdor. Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, 41(4) :778–820, 2006.
- [51] Charles T Clotfelter, Helen F Ladd, and Jacob L Vigdor. Teacher credentials and student achievement in high school a cross-subject analysis with student fixed effects. *Journal of Human Resources*, 45(3) :655–681, 2010.
- [52] Laurent Davezies, Xavier d’Haultfoeuille, and Denis Fougère. Identification of peer effects using group size variation. *The Econometrics Journal*, 12(3) :397–413, 2009.
- [53] Thomas Dee and James Wyckoff. Incentives, selection, and teacher performance : Evidence from impact. Technical report, National Bureau of Economic Research, 2013.
- [54] Cour des Comptes. *Rapport public thématique : Gérer les enseignants autrement, mai 2013*. Documentation française, 2013.

- [55] Weili Ding and Steven F Lehrer. Do peers affect student achievement in china's secondary schools? *The Review of Economics and Statistics*, 89(2) :300–312, 2007.
- [56] Esther Duflo, Pascaline Dupas, and Michael Kremer. Peer effects, teacher incentives, and the impact of tracking : Evidence from a randomized evaluation in kenya. *The American Economic Review*, 101 :1739–1774, 2011.
- [57] Esther Duflo, Rema Hanna, and Stephen P Rya. Incentives work : Getting teachers to come to school. *The American Economic Review*, 102(4) :1241–1278, 2012.
- [58] Dennis Epple and Richard Romano. Peer effects in education : A survey of the theory and evidence. *Handbook of social economics*, 1(11) :1053–1163, 2011.
- [59] Dennis Epple, Richard Romano, and Holger Sieg. Admission, tuition, and financial aid policies in the market for higher education. *Econometrica*, 74(4) :885–928, 2006.
- [60] Dennis Epple and Richard E Romano. Competition between private and public schools, vouchers, and peer-group effects. *American Economic Review*, pages 33–62, 1998.
- [61] Gabrielle Fack and Julien Grenet. When do better schools raise housing prices? evidence from paris public and private schools. *Journal of Public Economics*, 94(1) :59–77, 2010.
- [62] Marcel Fafchamps and Flore Gubert. The formation of risk sharing networks. *Journal of Development Economics*, 83(2) :326 – 350, 2007. Papers from a Symposium : The Social Dimensions of Microeconomic Behaviour in Low-Income Communities.
- [63] Clayton R Featherstone. Rank efficiency : Investigating a widespread ordinal welfare criterion. *Unpublished manuscript*, 2011.
- [64] Maria Marta Ferreyra. Estimating the effects of private school vouchers in multidistrict economies. *The American Economic Review*, pages 789–817, 2007.
- [65] David N Figlio. Boys named sue : Disruptive children and their peers. *Education Finance and Policy*, 2(4) :376–394, 2007.
- [66] David N Figlio and Lawrence W Kenny. Individual teacher incentives and student performance. *Journal of Public Economics*, 91(5) :901–914, 2007.
- [67] Edward B Fiske and Helen F Ladd. *When schools compete : A cautionary tale*. Brookings Institution Press, 2001.
- [68] Gigi Foster. It's not your peers, and it's not your friends : Some progress toward understanding the educational peer effect mechanism. *Journal of public Economics*, 90(8) :1455–1475, 2006.
- [69] Milton Friedman. *The role of government in education*. Rutgers University Press, 1955.

- [70] Milton Friedman. *Capitalism and freedom*. University of Chicago press, 2009.
- [71] Roland G Fryer. Teacher incentives and student achievement : Evidence from new york city public schools. *Journal of Labor Economics*, 31(2) :373–407, 2013.
- [72] David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *American Mathematical Monthly*, pages 9–15, 1962.
- [73] Michael S Garet, Stephanie Cronen, Marian Eaton, Anja Kurki, Meredith Ludwig, Wehmah Jones, Kazuaki Uekawa, Audrey Falk, Howard S Bloom, Fred Doolittle, et al. The impact of two professional development interventions on early reading instruction and achievement. ncee 2008-4030. *National Center for Education Evaluation and Regional Assistance*, 2008.
- [74] Michael S Garet, Andrew J Wayne, Fran Stancavage, James Taylor, Kirk Walters, Mengli Song, Seth Brown, Steven Hurlburt, Pei Zhu, Susan Sepanik, et al. Middle school mathematics professional development impact study : Findings after the first year of implementation. ncee 2010-4009. *National Center for Education Evaluation and Regional Assistance*, 2010.
- [75] Eric Gautier and Alexandre B Tsybakov. High-dimensional instrumental variables regression and confidence sets. *arXiv preprint arXiv :1105.2454*, 2011.
- [76] Stephen Gibbons, Stephen Machin, and Olmo Silva. Choice, competition, and pupil achievement. *Journal of the European Economic Association*, 6(4) :912–947, 2008.
- [77] Paul Glewwe, Nauman Ilias, and Michael Kremer. Teacher incentives. *American Economic Journal : Applied Economics*, 2(3) :205–27, 2010.
- [78] Dan Goldhaber and Emily Anthony. Can teacher quality be effectively assessed? national board certification as a signal of effective teaching. *The Review of Economics and Statistics*, 89(1) :134–150, 2007.
- [79] Paul Goldsmith-Pinkham and Guido W Imbens. Social networks and the identification of peer effects. *Journal of Business & Economic Statistics*, 31(3) :253–264, 2013.
- [80] Sarena F Goodman and Lesley J Turner. The design of teacher incentive pay and educational outcomes : Evidence from the new york city bonus program. *Journal of Labor Economics*, 31(2) :409–420, 2013.
- [81] Eric A Hanushek, John F Kain, Jacob M Markman, and Steven G Rivkin. Does peer ability affect student achievement? *Journal of applied econometrics*, 18(5) :527–544, 2003.
- [82] Douglas N Harris and Tim R Sass. Teacher training, teacher quality and student achievement. *Journal of public economics*, 95(7) :798–812, 2011.
- [83] Justine Hastings, Thomas Kane, and Douglas Staiger. Heterogeneous preferences and the efficacy of public school choice. *Unpublished manuscript*, 2008.

- [84] Justine S Hastings, Thomas J Kane, and Douglas O Staiger. Parental preferences and school competition : Evidence from a public school choice program. Technical report, National Bureau of Economic Research, 2005.
- [85] Justine S Hastings and Jeffrey M Weinstein. Information, school choice, and academic achievement : Evidence from two experiments\*. *The Quarterly journal of economics*, 123(4) :1373–1414, 2008.
- [86] John William Hatfield, Fuhito Kojima, and Yusuke Narita. Promoting school competition through school choice : a market design approach. *Unpublished manuscript*, 2011.
- [87] Yinghua He. Gaming the boston school choice mechanism in beijing. *Unpublished manuscript, Toulouse School of Economics*, 2012.
- [88] Victor Hiller and Olivier Tercieux. Choix d'écoles en france. *Revue économique*, 65(3) :619–656, 2014.
- [89] Peter Hinrichs. The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities. *Review of Economics and Statistics*, 94(3) :712–722, 2012.
- [90] Caroline M Hoxby. Does competition among public schools benefit students and taxpayers? *American Economic Review*, pages 1209–1238, 2000.
- [91] Caroline M Hoxby. *The economics of school choice*. University of Chicago Press, 2007.
- [92] Caroline Minter Hoxby. School choice and school productivity. could school choice be a tide that lifts all boats? In *The economics of school choice*, pages 287–342. University of Chicago Press, 2003.
- [93] Chang-Tai Hsieh and Peter J Klenow. Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics*, 124(4) :1403–1448, 2009.
- [94] Chang-Tai Hsieh and Miguel Urquiola. The effects of generalized school choice on achievement and stratification : Evidence from chile's voucher program. *Journal of public Economics*, 90(8) :1477–1503, 2006.
- [95] Chih-Sheng Hsieh and Lung-Fei Lee. A social interactions model with endogenous friendship formation and selectivity. Technical report, 2013.
- [96] Chih-Sheng Hsieh and Lung-Fei Lee. Specification and estimation of network formation and network interaction models with the exponential probability distribution. 2013.
- [97] Scott A Imberman, Adriana D Kugler, and Bruce I Sacerdote. Katrina's children : Evidence on the structure of peer effects from hurricane evacuees. *The American Economic Review*, 102(5) :2048–2082, 2012.
- [98] C Kirabo Jackson. Non-cognitive ability, test scores, and teacher quality : Evidence from 9th grade teachers in north carolina. Technical report, National Bureau of Economic Research, 2012.

- [99] Matthew O Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of economic theory*, 71(1) :44–74, 1996.
- [100] Brian A Jacob and Lars Lefgren. The impact of teacher training on student achievement quasi-experimental evidence from school reform efforts in chicago. *Journal of Human Resources*, 39(1) :50–79, 2004.
- [101] Brian A Jacob and Lars Lefgren. Can principals identify effective teachers? evidence on subjective performance evaluation in education. *Journal of Labor Economics*, 26(1) :101–136, 2008.
- [102] Brian A Jacob and Steven D Levitt. Rotten apples : An investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics*, 118(3) :843–877, 2003.
- [103] Michael C Jensen. Agency costs of free cash flow, corporate finance, and takeovers. *The American economic review*, pages 323–329, 1986.
- [104] Michael C Jensen and William H Meckling. Theory of the firm : Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4) :305–360, 1976.
- [105] Thomas J Kane, Stephanie K Riegg, and Douglas O Staiger. School quality, neighborhoods, and housing prices. *American Law and Economics Review*, 8(2) :183–212, 2006.
- [106] Thomas J Kane and Douglas O Staiger. Estimating teacher impacts on student achievement : An experimental evaluation. Technical report, National Bureau of Economic Research, 2008.
- [107] Changhui Kang. Classroom peer effects and academic achievement : Quasi-randomization evidence from south korea. *Journal of Urban Economics*, 61(3) :458–495, 2007.
- [108] Fuhito Kojima and Parag A Pathak. Incentives and stability in large two-sided matching markets. *The American Economic Review*, pages 608–627, 2009.
- [109] Victor Lavy. Evaluating the Effect of Teachers’ Group Performance Incentives on Pupil Achievement. *Journal of Political Economy*, 110(6) :1286–1317, December 2002.
- [110] Victor Lavy. Performance pay and teachers’ effort, productivity, and grading ethics. *American Economic Review*, 99(5) :1979–2011, 2009.
- [111] Victor Lavy, M Daniele Paserman, and Analia Schlosser. Inside the black box of ability peer effects : Evidence from variation in the proportion of low achievers in the classroom. *The Economic Journal*, 122(559) :208–237, 2012.
- [112] Victor Lavy and Edith Sand. The friends factor : how students’ social networks affect their academic achievement and well-being? Technical report, National Bureau of Economic Research, 2012.

- [113] Victor Lavy, Olmo Silva, and Felix Weinhardt. The good, the bad, and the average : evidence on ability peer effects in schools. *Journal of Labor Economics*, 30(2) :367–414, 2012.
- [114] Edward P Lazear. Educational production. *Quarterly Journal of Economics*, pages 777–803, 2001.
- [115] Edward P Lazear and Sherwin Rosen. Rank-order tournaments as optimum labor contracts. *The Journal of Political Economy*, pages 841–864, 1981.
- [116] Lung-Fei Lee. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, 72(6) :1899–1925, 2004.
- [117] Lung-fei Lee. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics*, 140(2) :333–374, 2007.
- [118] Audrey Light and Wayne Strayer. Determinants of college completion : School quality or student ability? *Journal of Human Resources*, pages 299–332, 2000.
- [119] Audrey Light and Wayne Strayer. From bakke to hopwood : Does race affect college attendance and completion? *Review of Economics and Statistics*, 84(1) :34–44, 2002.
- [120] Xu Lin. Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics*, 28(4) :825–860, 2010.
- [121] Linda Datcher Loury and David Garman. College selectivity and earnings. *Journal of labor Economics*, pages 289–308, 1995.
- [122] David S Lyle. Estimating and interpreting peer and role model effects from randomly assigned social groups at west point. *The Review of Economics and Statistics*, 89(2) :289–299, 2007.
- [123] W Bentley MacLeod and Miguel Urquiola. Anti-lemons : School reputation, relative diversity, and educational quality. 2012.
- [124] W Bentley MacLeod and Miguel Urquiola. Competition and educational productivity : Incentives writ large. 2012.
- [125] Elena Manresa. Estimating the structure of social interactions using panel data. *Unpublished Manuscript. CEMFI, Madrid*, 2013.
- [126] Charles F Manski. Identification of endogenous social effects : The reflection problem. *The review of economic studies*, 60(3) :531–542, 1993.
- [127] Pedro S Martins. Individual teacher incentives, student achievement and grade inflation. 2010.
- [128] Éric Maurin. *Le ghetto français : enquête sur le séparatisme social*. Seuil, 2004.
- [129] Edward Miguel and Michael Kremer. Worms : identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, 72(1) :159–217, 2004.

- [130] Jacob Mincer. Investment in human capital and personal income distribution. *The journal of political economy*, pages 281–302, 1958.
- [131] Karthik Muralidharan and Venkatesh Sundararaman. Teacher performance pay : Experimental evidence from india. *The Journal of Political Economy*, 119(1) :39–77, 2011.
- [132] Thomas J Nechyba. Mobility, targeting, and private-school vouchers. *American Economic Review*, pages 130–146, 2000.
- [133] Christopher Neilson. Targeted vouchers, competition among schools, and the academic achievement of poor students. 2013.
- [134] Parag A Pathak. The mechanism design approach to student assignment. *Annu. Rev. Econ.*, 3(1) :513–536, 2011.
- [135] Parag A Pathak and Tayfun Sönmez. Leveling the playing field : Sincere and sophisticated players in the boston mechanism. *The American Economic Review*, 98(4) :1636–1652, 2008.
- [136] Parag A Pathak and Tayfun Sönmez. School admissions reform in chicago and england : Comparing mechanisms by their vulnerability to manipulation. *American Economic Review*, 103(1) :80–106, 2013.
- [137] Marek Pycia. Stability and preference alignment in matching and coalition formation. *Econometrica*, 80(1) :323–362, 2012.
- [138] Steven G Rivkin, Eric A Hanushek, and John F Kain. Teachers, schools, and academic achievement. *Econometrica*, 73(2) :417–458, 2005.
- [139] Jonah E Rockoff. The impact of individual teachers on student achievement : Evidence from panel data. *American Economic Review*, pages 247–252, 2004.
- [140] Jonah E Rockoff and Benjamin B Lockwood. Stuck in the middle : Impacts of grade configuration in public schools. *Journal of Public Economics*, 94(11) :1051–1061, 2010.
- [141] Jonah E Rockoff and Cecilia Speroni. Subjective and objective evaluations of teacher effectiveness : Evidence from new york city. *Labour Economics*, 18(5) :687–696, 2011.
- [142] Jonah E. Rockoff, Douglas O. Staiger, Thomas J. Kane, and Eric S. Taylor. Information and employee evaluation : Evidence from a randomized intervention in public schools. *American Economic Review*, 102(7) :3184–3213, 2012.
- [143] Jon C Rogowski and Betsy Sinclair. Estimating the causal effects of social interaction with endogenous networks. *Political Analysis*, pages 1–13, 2012.
- [144] Matthew Ronfeldt, Susanna Loeb, and James Wyckoff. How teacher turnover harms student achievement. *American Educational Research Journal*, 50(1) :4–36, 2013.
- [145] Christiern Rose. Estimation of networks and spillover effects using panel data. *Unpublished Manuscript*, 2014.

- [146] Michael Rothschild and Lawrence J White. The analytics of the pricing of higher education and other services in which the customers are inputs. *Journal of Political Economy*, pages 573–586, 1995.
- [147] Jesse Rothstein. Does competition among public schools benefit students and taxpayers? comment. *The American Economic Review*, pages 2026–2037, 2007.
- [148] Jesse Rothstein. Teacher quality in educational production : Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, 125(1) :175–214, 2010.
- [149] Jesse Rothstein. Teacher quality policy when supply matters. Technical report, National Bureau of Economic Research, 2012.
- [150] Jesse Rothstein and Albert Yoon. Mismatch in law school. Technical report, National Bureau of Economic Research, 2008.
- [151] Jesse M Rothstein. Good principals or good peers? parental valuation of school characteristics, tiebout equilibrium, and the incentive effects of competition among jurisdictions. *The American Economic Review*, pages 1333–1350, 2006.
- [152] Bruce Sacerdote. Peer effects with random assignment : Results for dartmouth roommates. *The Quarterly Journal of Economics*, 116(2) :681–704, 2001.
- [153] Richard H Sander. A systemic analysis of affirmative action in american law schools. *Stanford Law Review*, pages 367–483, 2004.
- [154] Theodore W Schultz. Investment in human capital. *The American economic review*, pages 1–17, 1961.
- [155] Matthew G Springer, Dale Ballou, Laura Hamilton, Vi-Nhuan Le, JR Lockwood, Daniel F McCaffrey, Matthew Pepper, and Brian M Stecher. Teacher pay for performance : Experimental evidence from the project on incentives in teaching (point). *Society for Research on Educational Effectiveness*, 2011.
- [156] Douglas O Staiger and Jonah E Rockoff. Searching for effective teachers with imperfect information. *The Journal of Economic Perspectives*, 24(3) :97–117, 2010.
- [157] Ralph Stinebrickner and Todd R Stinebrickner. What can be learned about peer effects using college roommates? evidence from new survey data and students from disadvantaged backgrounds. *Journal of public Economics*, 90(8) :1435–1454, 2006.
- [158] Doug Williams. Do racial preferences affect minority learning in law schools? *Journal of Empirical Legal Studies*, 10(2) :171–195, 2013.
- [159] David J Zimmerman. Peer effects in academic outcomes : Evidence from a natural experiment. *Review of Economics and Statistics*, 85(1) :9–23, 2003.