



Inégalités scolaires, ségrégation et effets de pairs

Arnaud Riegert

► To cite this version:

Arnaud Riegert. Inégalités scolaires, ségrégation et effets de pairs. Education. EHESS, 2016. Français.
NNT: . tel-01333797

HAL Id: tel-01333797

<https://theses.hal.science/tel-01333797>

Submitted on 19 Jun 2016

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.

THÈSE DE DOCTORAT



En vue de l'obtention du grade de
Docteur en sciences économiques
de l'École des hautes études en sciences sociales

École doctorale Économie Panthéon-Sorbonne
Unité de recherche Paris-Jourdan Sciences Économiques

Inégalités scolaires, ségrégation et effets de pairs

Présentée par

Arnaud RIEGERT

et soutenue publiquement le

17 juin 2016

devant le jury composé de :

M. Éric MAURIN — Directeur de thèse

Mme Catherine MOISAN — Présidente du jury

M. Laurent DAVEZIES — Rapporteur

Mme Pascaline DUPAS — Rapporteuse

Remerciements

Je tiens en premier lieu à remercier Éric MAURIN qui a joué un rôle central dans toute ma formation en économie. Après avoir accepté de m'accueillir au master Politiques publiques et développement, en 2008, il a fortement encouragé la démarche d'évaluation du programme de tutorat Talens qui a donné lieu au chapitre 1 de cette thèse. Après quelques errements académiques, c'est encore lui qui m'a donné l'envie de venir terminer ma formation de recherche en économie et qui a accepté de diriger cette thèse. Il a été pour moi un directeur idéal, toujours disponible, de bon conseil tant sur les questions scientifiques que sur les nombreuses questions que je me suis posées sur mon parcours personnel. J'ai beaucoup appris auprès de lui et je lui suis grandement reconnaissant pour toute son aide.

J'ai la très grande chance de travailler avec Son Thierry Lx depuis dix ans sur des projets très variés, dont l'ensemble des chapitres de cette thèse. Au-delà de notre passion commune pour les sujets liés à l'éducation et la réduction des inégalités, nous avons su trouver un équilibre quasi parfait entre similarités et complémentarité, qui nous permet de nous comprendre rapidement et de nous répartir le travail efficacement. C'est un plaisir permanent de travailler avec Thierry, de profiter de sa créativité débordante et de son enthousiasme à toute épreuve. Je le remercie pour ces années de collaborations qui sont loin d'être terminées, mais aussi pour son amitié et son soutien dans les moments de doute.

Je remercie vivement Laurent DAVEZIES et Pascaline DUPAS pour leurs remarques et suggestions sur les travaux présentés dans cette thèse, en particulier sur le chapitre 1 dont ils ont été les premiers lecteurs. Merci à eux d’avoir accepté d’être les rapporteurs de cette thèse.

Tous les chapitres de cette thèse reposent sur des bases de données d’une qualité exceptionnelle produites par la Direction de l’évaluation, de la prospective et de la performance (Depp) au Ministère de l’Éducation nationale. Je tiens à remercier Catherine MOISAN, qui a permis et encouragé les collaborations entre la Depp et les chercheurs et qui a accepté d’être présidente de mon jury de thèse. Un grand merci également à Cédric AFSA, Olivier MONSO, Fabrice MURAT avec qui j’ai eu de nombreux échanges très enrichissants, à Caroline SIMONIS-SUEUR ainsi qu’à toutes les personnes qui constituent, nettoient et documentent en permanence les bases de données Scolarité et Océan, à la Depp et dans les service statistiques académiques.

Je remercie les très nombreux acteurs de l’éducation avec lesquels j’ai eu l’occasion d’échanger ces dernières années sur de nombreux sujets, qu’ils s’agisse des équipes pédagogiques et administratives des établissements, des associations d’éducation populaire, des membres de l’administration centrale et des rectorats, de l’ESENESR, de Canopé, des collectivités territoriales ou des cabinets ministériels. Je remercie particulièrement Anne-Christine LANG et l’unité lycées du Conseil régionale d’Île-de-France, Bénédicte ROBERT et l’équipe du DRDIE, ainsi qu’Olivier NOBLECOURT et le cabinet de Najat Vallaud-Belkacem.

J’ai également beaucoup collaboré avec le Conseil national d’évaluation du système scolaire, notamment dans le cadre du chapitre 3 de cette thèse. Je remercie très chaleureusement Nathalie MONS avec qui nous avons eu de nombreux échanges qui ont permis de beaucoup améliorer cet état des lieux de la ségrégation dans les collèges et lycées français, ainsi que toute son équipe de choc qui nous a aidé à produire et à diffuser cette étude à un large public : Sadiya BARKOUSS, Amandine BLANCHARD-SCHNEIDER, Thibault COUDROY, Arnaud GALINIÉ, Adeline GUBLER, Emily HELMEID et plus particulièrement Arthur HEIM. Ils ont fait et continuent à faire un travail remarquable que je tiens à saluer.

J’ai eu le plaisir d’effectuer toute ma formation en économie à l’École d’économie de Paris,

d’abord dans le master Politiques publiques et développement puis dans le programme doctoral Analyse et politique économiques, puis d’y enseigner depuis 2015. J’y ai découvert avec beaucoup d’intérêt les méthodes empiriques appliquées à l’économie grâce à des cours de très grande qualité. Je tiens à remercier l’ensemble des professeurs du master, et tout particulièrement Luc BEHAGHEL, Thomas BREDAS, Denis COGNEAU, Julien GRENET, Marc GURGAND, Sylvie LAMBERT, Thomas PIKETTY et bien entendu Éric MAURIN. Merci également à mes camarades de master et de doctorat, en particulier Camille TERRIER et Asma BENHENDA avec qui j’ai pu beaucoup échanger sur les sujets liés à l’éducation, ainsi qu’aux participants du séminaire d’économie appliquée. Je remercie enfin les équipes administratives pour leur efficacité et leur disponibilité, notamment Caroline BAUER, Marie FONT, Véronique GUILLOTIN, Béatrice HAVET, Damien HERPE, Weronika LEDUC, Pauline MARMIN, Marie-Christine PAOLETTI, Marie PHILIPPON, Sylvain RIFFÉ STERN et Anaëlle SIX, ainsi que France ARTOIS-MBAYE à l’EHESS.

Merci aux personnes avec qui j’ai eu la chance de travailler lors du stage très enrichissant que j’ai effectué au fonds d’expérimentation pour la jeunesse et qui m’ont beaucoup appris : Jean-Benoît DUJOL, Étienne GRASS, Malika KACIMI, Jérôme TEILLARD, Marie-Automne THÉPOT, Mathieu VALDENNAIRE et Augustin VICARD. Je remercie notamment Martin HIRSCH pour sa grande disponibilité, ses conseils et ses encouragements.

Je remercie également mes collègues de l’Insee et du Crest, en particulier Malik KOUBI et Corinne PROST pour la confiance qu’ils m’ont accordée dans le choix des sujets des études que j’ai menées à l’Insee, qui m’ont permis d’avancer simultanément sur ma thèse. Je les remercie notamment pour leurs suggestions qui ont permis de beaucoup améliorer le chapitre 4 de cette thèse. J’adresse également des remerciements particuliers à Xavier D’HAULTFŒUILLE qui m’a accueilli au laboratoire de microéconomie du Crest, à Roland RATHELOT avec qui nous avons pu échanger sur les aspects statistiques de la mesure de la ségrégation et à Alain TROGNON qui encadre la formation complémentaire décalée et m’a permis de terminer ma thèse tout en débutant un projet entrepreneurial.

Le chapitre 1 de cette thèse repose sur l’investissement des très nombreux acteurs du programme Talens de l’ENS. Je voudrais remercier ici l’équipe surmotivée qui le fait vivre depuis

10 ans : Claire SCOTTON, Son Thierry LY (encore lui !), Marie DARRASON, Gwenaël ROUDAUT, Muy-Cheng PEICH, Adeline PAUTRIZEL, Claire RIPAUT-BLANDIN, Pauline ROLLAND, Quentin LOUISIADE, Matthieu BAILLERGEAU, Tiphaine MALASEVIC ; les membres de la direction de l'ENS qui l'ont soutenu, notamment Véronique PROUVOST, Sophie FERMIGIER et Olivier ABILLON ; les générations de tuteurs, élèves, enseignants, CPE et proviseurs qui ont cru au projet. Merci aux nombreux tuteurs qui se sont investis dans les bureaux de l'association ou auprès du pôle PESU au-delà des séances de tutorat pour assurer sa pérennité, en particulier ces dernières années Marie HUET, Claire COLARD et « les chimistes » que je me permets de désigner collectivement.

Ceux qui me connaissent savent que je suis très facilement distrait par les nouveaux projets, et j'ai du compter sur un noyau dur de proches pour persévérer jusqu'au bout de cette thèse. Je voudrais donc remercier mes parents et mes amis qui m'ont soutenu et encouragé lors de toutes mes hésitations, ceux qui m'ont changé les idées et ceux qui ont entretenu ma passion pour les sujets liés à l'éducation et aux inégalités. Je pense notamment aux amis de Bibliothèques Sans Frontières qui font des choses exceptionnelles et que je suis heureux de retrouver régulièrement, et bien entendu à mes associés de Gryzz-Lab : Frédéric FOURCADE, Hai LY, Son Thierry LY (toujours lui !) et Adeline PAUTRIZEL, avec qui nous avons lancé une nouvelle aventure passionnante, et tous les collègues qui nous ont rejoint depuis.

Enfin, merci à Muy-Cheng pour tout ce qu'elle a fait pour moi et que je ne saurais énumérer.

Résumé

Cette thèse porte sur les enjeux de mixité dans l'institution scolaire en illustrant trois problématiques.

La première est celle de l'égalité des chances : comment promouvoir la mixité dans l'enseignement supérieur quand l'enseignement secondaire est lui-même ségrégué ? À l'aide d'une expérimentation aléatoire, nous évaluons l'impact d'un programme de tutorat visant à promouvoir l'accès aux filières d'excellence du supérieur pour des lycées issus de milieux populaires et montrons que ces dispositifs peuvent se heurter à des difficultés, notamment face aux élèves de niveau moyen, au risque de creuser les inégalités dans les lycées plus défavorisés.

La deuxième problématique est celle de l'influence des pairs : des modifications dans l'environnement social des élèves peuvent-elles avoir un effet sur la scolarité des jeunes ? Nous exploitons ici une expérience naturelle qui se produit en France lors de la transition collège-lycée, où nous montrons qu'un petit pourcentage d'élèves est affecté aléatoirement à sa classe de seconde. L'observation des trajectoires scolaires montre que les élèves fragiles qui retrouvent très peu de camarades de troisième dans leur classe de seconde sont plus susceptibles de redoubler et ont une probabilité moindre d'obtenir le baccalauréat.

La troisième problématique est celle de la ségrégation : dans quelle mesure les élèves fréquentent-ils des environnements sociaux différents en fonction de leurs caractéristiques personnelles ? Nous effectuons ici une analyse statistique détaillée qui permet de quantifier l'ampleur de la ségrégation sociale et scolaire, entre les établissements scolaires et entre les classes au sein de ceux-ci.

Mots-clés : éducation, inégalités, ségrégation, effets de pairs

Classification JEL : I24, D63

Abstract

This thesis examines three issues in relation with diversity in the educational system.

The first issue focuses on equal opportunities: how may we promote diversity in higher education when secondary education is segregated? We implemented a random experiment in order to evaluate the impact of a tutoring program that aims at encouraging underprivileged high school students to enroll into the most selective tracks of higher education. We show that such programs may face difficulties, especially when dealing with mid-level students, which may lead to a risk of increased inequality in underprivileged high schools.

The second issue is the influence of the peer group: may modifications in the social environment of high school students have an effect on their subsequent education? We exploit a natural experiment occurring in France whereby a small fraction of 10th grade students are randomly allocated to their classes. We find that the most fragile students who end up in classes where they meet very few former classmates from 9th grade are more enclined to repeat the grade and have a lower probability of graduating high school.

The third issue is segregation: to what extent do students experience different social environments in their classes or schools depending on their individual characteristics? We provide a detailed statistical analysis which allows us to measure quantitatively the extent of social and academic segregation between schools and between the classes within those schools.

Keywords: education, inequality, segregation, peer effects

JEL classification: I24, D63

Table des matières

Remerciements	i
Résumé et mots-clefs	vii
Abstract and keywords	ix
Table des matières	xi
Table des figures	xvii
Liste des tableaux	xxi
Introduction	1
L'évaluation du programme Talens	2
Le rôle des anciens camarades de classe	6
L'état des lieux de la ségrégation	9
La mesure de la ségrégation dans la durée	14
1 Elite Tutors and Underprivileged High-School Students	19
avec Son Thierry LY et Éric MAURIN	
1.1 Introduction	19
1.2 The experiment	23

1.2.1	Institutional context	23
1.2.2	The program and its objectives	24
1.2.3	Identification of volunteers	25
1.2.4	Random selection of eligible students	26
1.2.5	Program content, tutors and take up	29
1.3	Data	32
1.4	Effects on achievement and choices	33
1.4.1	Effects on high school achievement	33
1.4.2	Effects on access to selective undergraduate programs	36
1.5	Mechanisms: the role of tutors and peers	39
1.5.1	Tutors' characteristics	39
1.5.2	Peer group influence	40
1.6	Discussion	42
1.7	Bibliography	44
1.A	Additional tables	46
2	Persistent Classmates: How Familiarity with Peers Protects from Disruptive School Transitions	55
	avec Son Thierry Ly	
2.1	Introduction	56
2.2	Institutional context and data	60
2.2.1	The high school curriculum in France	60
2.2.2	The class-assignment mechanism	64
2.2.3	Data	66
2.3	Identification	68
2.3.1	Definition of similar-file groups	71
2.3.2	Empirical evidence of random assignment	73
2.3.3	Description of the SF sample	79
2.4	Results	79
2.4.1	Freshman-year class characteristics and achievement	79
2.4.2	The protective role of familiarity with classmates	83
2.5	Robustness checks	89

2.5.1	Alternative SF group specifications	89
2.5.2	Estimation based on the impact of SF student allocation on their class-mates	91
2.6	Discussion and conclusion	97
2.7	Bibliography	99
2.A	Additional tables	101
3	La ségrégation sociale et scolaire en France : un état des lieux quantitatif	107
	avec Son Thierry LY	
3.1	Introduction	107
3.2	Méthodologie	111
3.2.1	Comment mesurer la ségrégation ?	111
3.2.2	L'indice d'exposition	116
3.2.3	Les données mobilisées	117
3.3	La ségrégation sociale et scolaire entre les établissements	118
3.3.1	Au niveau national	118
3.3.2	Disparités géographiques	122
3.4	La ségrégation sociale et scolaire entre les classes de chaque établissement	125
3.4.1	Analyse macroscopique	128
3.4.2	Mieux mesurer la ségrégation intra-établissement	131
3.4.3	Le rôle des langues et des options	136
3.5	Évolution de la ségrégation	141
	Conclusion	146
3.6	Bibliographie	147
3.A	Les indices de ségrégation	149
3.A.1	Définitions	149
3.A.2	Discussion	151
3.B	Mesures alternatives de la ségrégation	153
4	Measuring Social Environment Mobility	159
	avec Son Thierry LY	
4.1	Introduction	159

4.2	Literature review	162
4.2.1	Measuring income inequality	162
4.2.2	The equalizing force of income mobility	164
4.2.3	Measuring segregation	169
4.3	Social environment mobility	173
4.3.1	Breaking out of the clusters	173
4.3.2	The limited equalizing force of social environment mobility	178
4.3.3	Dealing with attrition	181
4.3.4	Nested levels of clustering	184
4.4	Case study: SEM in French middle schools	187
4.5	Bibliography	190
4.A	Mathematical proofs	192
4.A.1	Equivalence of the two definitions of the Gini index	192
4.A.2	Normalization of the exposure index	193
4.A.3	Normalized exposure is equal to the exposure gap	194
4.A.4	Upper envelope of the Lorenz curve of lifetime social environments . . .	196

Table des figures

2.1	Composition of the typical classroom from a non-repeating student's point of view	63
2.2	Class assignment of similar-file students	70
3.1	Ségrégation sociale et scolaire inter-établissements au niveau national	119
3.2	Carte de la ségrégation sociale inter-établissements en classe de troisième . . .	124
3.3	Carte de la ségrégation scolaire inter-établissements en classe de troisième . .	125
3.4	Ségrégation scolaire entre établissements au niveau troisième par académie . .	126
3.5	Corrélation entre la ségrégation sociale et la ségrégation scolaire	127
3.6	Ségrégation sociale et scolaire inter- et intra-établissement au niveau national .	129
3.7	Carte de la ségrégation sociale totale en classe de troisième	132
3.8	Carte de la ségrégation scolaire totale en classe de troisième	133
3.9	Ségrégation scolaire entre classes au niveau troisième par académie	134
3.10	Évolution de la ségrégation en sixième	142
3.11	Évolution de la ségrégation en troisième	143
3.12	Évolution de la ségrégation en seconde	144
3.B.1	Indice d'exposition mesuré sur les élèves CSP- ou en retard scolaire	154
3.B.2	Indice de dissimilarité	155
3.B.3	Indice de Theil	156

4.1	The Lorenz curve for income inequality	163
4.2	A social network graph with interactions in clusters	171
4.3	Segregation between upper class (full circles) and middle or lower class (plain circles) students.	173
4.4	Social environment mobility in the case of school segregation	174
4.5	The Lorenz curve for social environments	177
4.6	The upper envelope of lifetime social environment Lorenz curves	180
4.7	Attrition	182
4.8	Normalized exposure in French middle schools, cohort 2009	188

Liste des tableaux

1.1	The characteristics of students who volunteer to participate in the program . .	26
1.2	Take up rates, by cohort and ability group	30
1.3	The effect of the treatment on performance on high school exit examinations .	34
1.4	The effect of the treatment on access to selective undergraduate programs (CPGE)	38
1.5	The effect of tutor and assigned peers characteristics on grade 11 achievement	41
1.A.1	Balancing test on the baseline survey	47
1.A.2	Balancing test on the sample of volunteer students	48
1.A.3	The effect of the treatment on performance on high school exit examinations, by cohort	49
1.A.4	The effect of the treatment on performance on high school exit examinations, by track	50
1.A.5	The effect of the treatment on scores at the <i>Baccalauréat</i> in the scientific track .	51
1.A.6	The effect of the treatment on scores at the <i>Baccalauréat</i> in the non scientific tracks	52
2.1	Descriptive statistics on students' characteristics	69
2.2	Student's class characteristics regressed on own anonymous exam score: Evi- dence of the random assignment of similar-file students	75

2.3	Student's class characteristics regressed on behavior score: Evidence of the random assignment of similar-file students	78
2.4	Effect of class characteristics on high school outcomes	81
2.5	Effect of persistent classmates on high school outcomes with and without controlling for other class characteristics	84
2.6	Distribution of the PC effect	86
2.7	Which peers do matter? Decomposition of the PC effect on students at risk. .	90
2.8	Robustness check: effect of PC on low-ability students' repetition rate using different specifications of the SF fixed effect	92
2.9	IV exogeneity test	95
2.10	Effect of PC on high school outcomes using the IV strategy	96
2.A.1	Student's class characteristics regressed on own anonymous exam score : Evidence of the random assignment of similar-file students	102
2.A.2	Anonymous exam score regressed on class characteristics : Evidence of the random assignment of similar-file students	103
2.A.3	Raw correlation between class characteristics and high school outcomes	104
2.A.4	Effect of class characteristics on high school outcomes for the sample "at risk"	105
3.1	Classification des PCS de la DEPP	114

Introduction

Cette thèse porte sur les enjeux de mixité dans l'institution scolaire. Elle illustre trois problématiques qui se posent dans une démarche de recherche de mixité sociale ou scolaire à l'école.

La première problématique est celle de la promotion de l'égalité des chances : comment peut-on promouvoir la mixité dans l'enseignement supérieur quand l'enseignement secondaire est lui-même ségrégué ? À l'aide d'une expérimentation par assignation aléatoire, nous évaluons l'impact d'un programme de tutorat visant à promouvoir l'accès aux filières d'excellence du supérieur pour des lycées issus de milieux populaires et montrons que ces dispositifs peuvent se heurter à des difficultés, notamment face aux élèves de niveau moyen, au risque de creuser les inégalités dans les lycées plus défavorisés.

La deuxième problématique est celle de l'influence des pairs : des modifications dans l'environnement social des élèves peuvent-elles avoir un effet sur la scolarité des jeunes, même à un stade avancé ? Nous exploitons ici une expérience naturelle qui se produit en France lors de la transition collège-lycée, où nous montrons qu'un petit pourcentage d'élèves est affecté aléatoirement à sa classe de seconde. L'observation des trajectoires scolaires montre notamment que les élèves fragiles qui retrouvent très peu de camarades de troisième dans leur classe de seconde sont plus susceptibles de redoubler cette classe et ont une probabilité moindre d'obtenir le baccalauréat.

La troisième problématique est celle de la ségrégation : dans quelle mesure les élèves fréquentent-ils des environnements sociaux différents, dans leur classe ou dans leur établissement, en fonction de leurs caractéristiques personnelles ? Nous effectuons ici une analyse statistique détaillée qui permet de quantifier l'ampleur de la ségrégation sociale et scolaire, entre les établissements scolaires et entre les classes au sein de ceux-ci. Nous introduisons également un outil qui permet de mesurer la diversité des environnements sociaux fréquentés par un élève au cours de sa scolarité.

Ces sujets sont inspirés par des expériences de terrain et des rencontres avec différents acteurs de l'Éducation nationale. Ils ont donné lieu aux quatre chapitres de cette thèse, dont le contenu est difficilement accessible aux lecteurs non-spécialistes en économétrie et en statistiques. Les résumés ci-dessous ont pour but d'en expliquer les résultats importants et les grands principes de la méthodologie employée, en évitant autant que possible une trop grande technicité.

L'évaluation du programme Talens

Le chapitre 1 évalue l'impact du programme de tutorat Talens, créé à l'École normale supérieure en 2006 sous la forme d'une association étudiante et repris par la direction de l'École en 2010 afin de le pérenniser. Au moment de cette reprise, le principe a été posé d'évaluer rigoureusement son impact sur les bénéficiaires du programme, notamment en termes de résultats scolaires et d'orientation.

Cette évaluation a été rendue possible par un financement du fonds d'expérimentation pour la jeunesse, qui a permis d'augmenter temporairement la capacité du dispositif pour accueillir 220 élèves en septembre 2010, afin d'atteindre une taille d'échantillon suffisante pour permettre les analyses statistiques. Elle était motivée par la forte montée en charge, en 2008, des dispositifs de type « cordées de la réussite » qui mettent en relation des élèves des formations sélectives du supérieur avec des lycéens et collégiens de milieux populaires pour favoriser leur accès à l'enseignement supérieur et notamment aux filières d'excellence. Outre l'investissement bénévole des tuteurs des « têtes de cordées », ces dispositifs reposent essentiellement sur des financements publics, et la question de leur efficacité se pose, notamment en comparaison à d'autres dispositifs

d'accompagnement. Or, il n'existait aucune évaluation quantitative rigoureuse de l'impact de ces dispositifs.

La méthode utilisée pour cette évaluation est celle de l'assignation aléatoire, qui consiste à tirer au sort les bénéficiaires du programme parmi les volontaires, pour comparer ensuite les résultats des élèves sélectionnés (groupe « test » ou « de traitement ») et non-sélectionnés (groupe « témoin » ou « de contrôle »). Cette méthode, souvent utilisée pour les tests cliniques et popularisée en économie – notamment du développement – par Esther Duflo, permet de s'assurer que la comparaison des résultats (notes obtenues, orientation, etc.) porte sur des groupes comparables : si la taille d'échantillon est suffisante, les deux groupes auront des compositions similaires dans toutes les dimensions, qu'elles soient observables ou non. Par exemple, nous avons pu vérifier que les notes du brevet et de seconde, les professions des parents et d'autres caractéristiques étaient bien équilibrées entre les deux groupes ; mais la méthode de tirage au sort permet aussi de s'assurer que les groupes sont similaires, en moyenne, sur des caractéristiques non observées telles que la motivation ou la curiosité. Ces caractéristiques non observées sont susceptibles d'être liées aux résultats scolaires et il est donc important que les deux groupes comparés soient similaires sur ces points. Il s'agit là d'un avantage important de la méthode de l'assignation aléatoire par rapport aux méthodes où un groupe de contrôle est constitué à partir d'élèves similaires selon des caractéristiques observables et mesurables uniquement.

Cette méthode est encore peu utilisée, notamment en France, pour évaluer l'impact d'un dispositif expérimental ou d'une politique publique, car elle peut impliquer une logistique complexe et peut poser des questions éthiques de rupture d'égalité. Une des missions du fonds d'expérimentation pour la jeunesse était précisément de faire la pédagogie de cet outil d'évaluation et de diffuser les bonnes pratiques. Le principe fondamental est que cette rupture d'égalité entre les deux groupes doit être temporaire et ne doit pas priver une partie de l'échantillon des dispositifs de droit commun. Une assignation aléatoire est particulièrement adaptée lorsque le dispositif à évaluer fait face à un excès de demande et qu'une sélection est nécessaire. En ce qui concerne le programme Talens, le nombre de volontaires a systématiquement dépassé la capacité d'accueil du programme ; avant l'expérimentation, la sélection se faisait en lien avec les enseignants et sur la base d'entretiens pour mesurer la motivation des élèves. Cette sélection avait cependant un

côté arbitraire et parfois difficile à expliquer aux élèves non retenus. Pour ces raisons, le principe du tirage au sort a été globalement bien accepté par les lycées partenaires.

Le programme Talens commence à la fin de l'été qui précède l'entrée en première, avec une semaine de « campus » à l'École normale supérieure à laquelle tous les élèves sélectionnés sont conviés. Les tuteurs volontaires organisent des activités pour préparer la rentrée scolaire et acquérir des méthodes de travail efficaces et des animations ont lieu le soir pour renforcer la cohésion du groupe. En septembre, les élèves choisissent un thème de tutorat (sciences, sciences humaines ou littérature) et sont affectés à des groupes de tutorat. Chaque groupe contient six lycéens et un tuteur de l'ENS, et ils se retrouvent toutes les deux à trois semaines pour des séances de tutorat dont les objectifs sont de faire découvrir le fonctionnement de l'enseignement supérieur et de les préparer aux attentes académiques des différentes filières. Cet accompagnement dure pendant toute l'année de première et, pour ceux qui souhaitent poursuivre, toute l'année de terminale.

Nous avons comparé les résultats obtenus au baccalauréat par les bénéficiaires du programme avec ceux des élèves qui s'étaient portés volontaires mais n'avaient pas été retenus par le tirage au sort. La comparaison montre une différence de moyennes proche de zéro, c'est-à-dire que l'impact du programme sur les notes obtenues au baccalauréat, que ce soit en fin de première ou sur la note finale de terminale n'est pas statistiquement significatif. En d'autres termes, le programme n'a pas contribué à augmenter les résultats des élèves bénéficiaires au baccalauréat.

Un second objectif du programme était d'augmenter les ambitions des lycéens et de les encourager à postuler à des formations sélectives. Nous avons donc comparé les taux d'accès aux Classes Préparatoires aux Grandes Écoles (CPGE), formations de deux ans qui mènent aux écoles les plus sélectives et par lesquelles sont passés la plupart des tuteurs de l'ENS. Là encore, la différence entre les deux groupes est non significative, c'est-à-dire que le programme n'a pas eu d'impact sur le taux d'accès en CPGE. La différence reste nulle pour le taux d'accès en deuxième année de CPGE, c'est-à-dire qu'il n'y a pas non plus d'impact sur le risque d'abandon en cours de première année de CPGE.

Ces résultats peuvent être surprenants : des élèves motivés et volontaires qui passent une

partie de leurs samedis après-midis à suivre des séances de tutorat auprès d'élèves de l'ENS ne voient pas leurs notes progresser et ne s'orientent pas plus vers les CPGE que leurs camarades qui n'ont pas pu bénéficier du programme. Pour mieux les comprendre, nous avons reproduit l'analyse pour deux sous-groupes de l'échantillon, en fonction des notes obtenues en seconde. Nous constatons alors que l'impact du programme de tutorat a été très positif sur la moitié des élèves ayant obtenu les moyennes générales de seconde les plus élevées, alors qu'il a été très négatif sur la moitié des élèves ayant obtenu les moyennes générales de seconde les plus basses. Au final, les inégalités entre ces deux groupes ont été creusées par le programme Talens, qui a eu un effet de stimulation sur les élèves les plus solides et un effet déprimant sur les plus fragiles.

Nous avons conduit en 2016 une enquête auprès des anciens bénéficiaires pour tenter d'éclairer les raisons de cet effet négatif. La première hypothèse était que le programme a un effet décourageant sur les élèves les plus fragiles, en leur faisant prendre conscience des difficultés qui les attend dans l'enseignement supérieur ; cette hypothèse n'est pas confirmée par l'enquête, qui montre que les élèves ont pour la plupart trouvé leurs tuteurs compréhensibles et pédagogues, y compris parmi les élèves les plus fragiles. La deuxième hypothèse était que le programme prenait trop de temps aux élèves, qui avaient en conséquence moins de temps pour réviser leurs cours et effectuer leurs devoirs. Le programme était en effet coûteux en temps, si on cumule la durée des séances elles-mêmes, les temps de transport et le temps passé avec les amis d'autres lycées rencontrés dans le cadre du tutorat. L'enquête confirme que beaucoup d'élèves, notamment parmi les plus fragiles, estimaient que le programme prenait beaucoup de temps et avait une incidence négative sur leur travail scolaire (le contenu des séances de tutorat n'étant pas directement lié au programme de lycée).

Les résultats de cette étude doivent permettre d'ouvrir le débat sur l'efficacité des « cordées de la réussite ». L'évaluation montre que des élèves peuvent se porter volontaires pour une préparation à l'enseignement supérieur et être impactés négativement sur leurs résultats scolaires. L'enquête menée sur les anciens bénéficiaires suggère par ailleurs que cet effet négatif n'est pas uniquement le fait du contenu particulièrement académique du programme Talens en comparaison avec les autres cordées. La réflexion doit donc s'ouvrir parmi les établissements porteurs de tels projets de leur impact sur chaque profil d'élève, ainsi qu'aux implications sur des sujets

comme la sélection des élèves, le rythme et le contenu du programme ou encore la formation des tuteurs.

Le rôle des anciens camarades de classe

Dans le chapitre 2, nous nous intéressons à l'influence de la composition de la classe de seconde sur la suite du parcours scolaire au lycée. La transition collège-lycée est un moment difficile de l'éducation secondaire : la troisième et la seconde sont les classes les plus redoublées et la seconde est la dernière année de scolarité obligatoire pour les élèves n'ayant pas d'année de retard. Cette transition se traduit également par un nouvel environnement où les attentes des professeurs sont plus élevées, les élèves commencent à préparer leur entrée dans l'enseignement supérieur, et ils doivent, le plus souvent, changer d'établissement. Ce changement d'établissement a pour effet de modifier profondément les réseaux sociaux des élèves : en moyenne, un élève de seconde retrouve dans son lycée huit camarades de sa classe de troisième. Parmi eux, 1,7 élèves en moyenne sont dans la même classe de seconde : les élèves venant d'une même classe de troisième sont le plus souvent séparés entre les classes de seconde par les proviseurs, notamment par crainte de problèmes de discipline. Il s'agit du seul moment de la scolarité secondaire où les élèves connaissent un tel changement.

Nous avons souhaité savoir dans quelle mesure ce changement brutal de composition des classes pouvait expliquer les difficultés rencontrées par les élèves de seconde. Pour cela, nous avons pu nous appuyer sur une spécificité de la transition entre la troisième et la seconde : il s'agit du seul moment de la scolarité secondaire où les équipes pédagogiques ne connaissent pas les élèves au moment de la constitution des classes. La répartition des élèves entre les classes de seconde, généralement effectuée au tout début de l'été par le proviseur, éventuellement avec l'aide des CPE et de certains professeurs, se fait sur la seule base des dossiers scolaires remis lors de l'inscription administrative. Il arrive alors que certains dossiers soient extrêmement similaires « sur le papier », correspondant à deux élèves venant du même collège et de la même classe de troisième, ayant choisi les mêmes langues vivantes et options et ayant obtenu des notes très proches à la fois dans les matières scientifiques et dans les matières littéraires, et étant de plus du même sexe, du même âge et de la même origine sociale.

Cette situation est très rare : elle concerne environ 2% des dossiers. Considérons deux élèves dans cette situation – appelons-les Aurélien et Benoît –, venant de la même classe de troisième d'un même collège et ayant des profils quasi identiques sur le papier. Deux options s'offrent alors au proviseur : il peut les regrouper dans la même classe de seconde, ou les affecter à deux classes différentes, par exemple la seconde A et la seconde B. Cette décision peut être motivée par différents critères, par exemple le souhait ou non de regrouper dans une même classe les élèves ayant choisi une certaine option. Dans notre étude, nous nous concentrons uniquement sur les situations où le proviseur fait le choix de séparer ces élèves similaires. Il a alors, là encore, deux possibilités : affecter Aurélien à la seconde A et Benoît à la seconde B, ou affecter Aurélien à la seconde B et Benoît à la seconde A. Pour prendre cette décision, le proviseur ne dispose d'aucune information autre que celles contenues dans les dossiers d'inscription, qui sont quasiment identiques pour les deux élèves. Pour départager les deux élèves, il ne reste alors plus que le hasard. Ainsi, par le fruit du hasard, deux élèves quasi identiques sur le papier commencent leur scolarité au lycée dans des classes de seconde différentes.

Il est important de noter que cette affectation aléatoire a un caractère exceptionnel : le plus souvent, le profil de l'élève a une influence sur les caractéristiques de la classe de seconde à laquelle il sera affecté. Il n'est alors pas possible, en général, d'isoler l'effet des caractéristiques de la classe sur la suite de la scolarité de l'élève, qui peut être affectée par de nombreux facteurs. Dans notre cas cependant, si nous parvenons à identifier un nombre suffisamment important de scénarios comme celui décrit ci-dessus, nous pourrions constituer des échantillons d'élèves aux profils quasi identiques mais ayant connu *par le fruit du hasard* des environnements de classe de seconde différents. Il devient alors possible de mesurer l'effet causal de certaines caractéristiques de classe sur la réussite scolaire des élèves.

En utilisant les bases de données du Ministère de l'Éducation nationale, qui contiennent la quasi totalité des informations présentes dans les fiches d'inscription dont disposent les proviseurs, nous avons identifiés près de 30 000 élèves dans cette situation. Nous avons alors pu comparer les parcours scolaires des lycéens selon que le hasard les avait affecté à une classe ayant plus ou moins de bons élèves, plus ou moins de filles, plus ou moins d'élèves issus de milieux favorisés, à une classe plus ou moins grande, ou à une classe dans laquelle il retrouvaient plus ou

moins de camarades de leur classe de troisième. Notre analyse montre des effets très faibles sur la scolarité de la plupart des caractéristiques énoncées ci-dessus, à l'exception de la dernière : un élève qui, par le fruit du hasard, retrouve plus de ses camarades de classe de troisième en seconde, aura une probabilité moins importante de redoubler la classe de seconde et une probabilité plus importante d'obtenir le baccalauréat.

Ces résultats ont une portée importante. Les nombreux travaux quantitatifs portant sur les effets de la composition de classe peinent à se mettre d'accord sur des résultats solides. Surtout, ces résultats ne permettent pas, pour l'heure, de formuler des recommandations précises sur la constitution des classes : si, par exemple, la présence de filles a un effet bénéfique sur la scolarité, le fait de déplacer des filles d'une classe à une autre aura un effet positif sur les uns et un effet négatif sur les autres. Il faut alors compter sur des effets différenciés ou non-linéaires pour espérer pouvoir optimiser les compositions de classe. Le cas des anciens camarades de classe est différent : si on considère deux élèves d'une même classe de troisième, chacun isolé dans une classe de seconde, le fait de les regrouper aura un effet positif sur ces deux élèves sans impacter négativement les autres élèves qu'ils ne connaissaient pas.

L'objet de cette étude n'est pas d'affirmer que les proviseurs doivent regrouper dans une même classe de seconde tous les élèves issus d'une même classe de troisième. Tout d'abord, les données dont nous disposons ne permettent d'observer que des classes où le nombre d'anciens camarades de classes retrouvés en seconde est compris entre zéro et quatre (il existe trop peu de cas où ce nombre est supérieur pour tirer des conclusions statistiquement valables). Il s'agit donc avant tout de veiller à éviter les situations d'isolement complet, en s'assurant dans la mesure du possible que chaque élève retrouve quelques anciens camarades de troisième dans sa nouvelle classe de seconde. Regrouper dans une même classe de seconde un grand nombre d'élèves d'une même classe de troisième est par ailleurs susceptible d'avoir pour effet d'augmenter la ségrégation sociale et scolaire entre les classes. Nous observons par ailleurs que les élèves les plus sensibles à la composition de la classe de seconde sont les élèves les plus fragiles scolairement et issus des classes populaires. La transition vers le lycée est en effet particulièrement difficile pour ces élèves ; à l'inverse, les élèves les plus solides ou issus des classes aisées sont peu impactés par le nombre de camarades retrouvés. L'effort doit donc se porter avant tout sur ces élèves fragiles. Enfin, la

recommandation issue de cette étude est de nature statistique, et ne doit pas supplanter des recommandations individuelles qui pourraient avoir été formulées par l'équipe pédagogique du collège d'origine (par exemple, de ne pas regrouper deux élèves jugés perturbateurs lorsqu'ils sont ensemble).

Un deuxième aspect de cette étude rend notre résultat simple à implémenter : le résultat porte sur le nombre de *camarades de classe* de troisième retrouvés, et non sur le nombre d'*amis* retrouvés. Il n'est donc pas nécessaire de mettre en place une logistique complexe pour recueillir les souhaits de regroupement formulés par les élèves, alors même que l'exercice de composition des classes est déjà particulièrement difficile. Nous montrons que les élèves sont autant impactés par le fait de retrouver des camarades du même sexe ou non, ou des camarades d'un niveau scolaire similaire ou non. Ces résultats indiquent qu'il s'agit avant tout d'un effet de *familiarité* avec le nouvel environnement de seconde : le fait de retrouver quelques visages connus a un effet rassurant à un moment charnière de la scolarité.

L'état des lieux de la ségrégation dans les collèges et lycées français

Le chapitre 3 a pour objet de dresser un portrait de la composition des collèges et des lycées français. Cette étude permet de décrire quantitativement l'ampleur de la ségrégation sociale et scolaire dans l'enseignement secondaire, un phénomène largement discuté dans le débat public mais sur lequel des informations chiffrées manquaient. En effet, il s'agit à notre connaissance du premier état de lieux exhaustif de la ségrégation sociale et scolaire dans les collèges et lycées français, portant à la fois sur les dimensions inter- et intra-établissement. L'étude met en évidence une forte différence entre les environnements sociaux fréquentés par les collégiens et lycéens français en fonction de leur origine sociale ou de leur niveau scolaire. Par exemple, un élève issu d'un milieu aisé compte en moyenne dans sa classe deux fois plus d'élèves également issus d'un milieu aisé qu'un élève issu d'un milieu populaire ou moyen. De même, les meilleurs élèves comptent en moyenne dans sa classe deux fois plus d'élèves également d'un très bon niveau scolaire que les élèves de niveau faible ou moyen. Ces chiffres sont inquiétant à deux titres : les

différences d'environnement en fonction de l'origine sociale ou du niveau scolaire sont susceptibles d'aggraver les inégalités scolaires ; de plus, cet entre-soi est un obstacle à l'apprentissage de la citoyenneté et du vivre-ensemble.

Afin de quantifier ces différences d'environnement fréquentés par les élèves en fonction de leurs caractéristiques, nous utilisons un indice de ségrégation appelé l'indice d'exposition. Cette indice est compris entre 0 et 100% : lorsqu'il vaut zéro, tous les élèves ont une composition de classe ou d'établissement similaire quelle que soit leur origine ; lorsqu'il vaut 100%, les élèves sont complètement isolés en fonction de leurs caractéristiques (par exemple, chaque classe ne contient que des élèves issus des milieux aisés ou n'en contient aucun). L'indice est défini par la différence entre la proportion de camarades de classe appartenant à une certaine catégorie (issus des milieux aisés ou ayant un niveau scolaire élevé) selon qu'on appartient ou non, soi-même, à cette catégorie. Nous mesurons ici deux formes de ségrégation :

La ségrégation sociale mesure la différence entre la proportion d'élèves issus des milieux aisés, ou « CSP+ »¹, dans l'environnement d'un élève (sa classe ou son établissement scolaire) selon qu'il est lui-même CSP+ ou non. La proportion d'élèves CSP+ dans la population est d'environ 20% : en l'absence de ségrégation sociale, chaque élève devrait donc compter dans sa classe 20% d'élèves CSP+. Lorsqu'il y a ségrégation, les élèves eux-mêmes CSP+ compteront en moyenne plus de CSP+ dans leur classe que les élèves non-CSP+. Si par exemple ils en comptent 36% alors que les non-CSP+ n'en comptent que 16%, l'indice de ségrégation vaudra 20% ($36 - 20$).

La ségrégation scolaire mesure la différence entre la proportion d'élèves ayant obtenu les meilleures notes au diplôme national du brevet (on considère les élèves du premier quart de la distribution des notes, que l'on appellera les « bons élèves » par souci de simplicité) dans l'environnement d'un élève (sa classe ou son établissement scolaire) selon qu'il fait lui même partie de ce groupe ou non. En l'absence de ségrégation scolaire, chaque élève devrait compter dans sa classe 25% de bons élèves. Lorsqu'il y a ségrégation, ces élèves compteront en moyenne plus d'élèves faisant également partie de ce groupe dans leur classe que les élèves moyens ou faibles. Si par exemple ils en comptent 55% alors que les autres élèves n'en compte que 15%, l'indice de ségrégation vaudra 40% ($55 - 15$).

1. CSP : catégorie socio-professionnelle.

Les définitions ci-dessus mentionnent des proportions d'élèves dans son « environnement » d'établissement ou de classe. La proportion d'élèves CSP+ dans l'environnement d'établissement d'un élève fait référence au nombre d'élèves CSP+ présents dans l'ensemble des classes de l'établissement d'un niveau donné (par exemple, toutes les classes de quatrième), divisé par le nombre total d'élèves de ce niveau. À l'inverse, la proportion d'élèves CSP+ dans l'environnement de classe d'un élève fait référence au nombre d'élèves CSP+ présents dans sa classe (par exemple, la quatrième A). Lorsqu'on effectue la mesure en utilisant les environnements d'établissement, le chiffre obtenu correspond à la ségrégation *inter-établissements*, c'est-à-dire qu'il mesure à quel point les collèges et lycées accueillent des élèves différents. Lorsqu'on effectue la mesure en utilisant les environnements de classe, le chiffre obtenu correspond à la ségrégation totale entre classes, c'est-à-dire à quel point les classes de tous les établissements accueillent des élèves différents. En prenant la différence entre les deux chiffres, on obtient la valeur de la ségrégation *intra-établissement*, c'est-à-dire à quel point les classes d'un établissement donné accueillent des élèves différents.

Notre premier constat est celui d'une forte ségrégation sociale entre les établissements : l'indice de ségrégation sociale inter-établissements varie en effet entre 16 et 19% selon le niveau considéré. Ainsi, alors que la proportion de CSP+ parmi les collégiens est d'environ 22%, les élèves eux-mêmes CSP+ en comptent en moyenne 34% dans leur établissement alors que les autres n'en comptent que 18%, soit environ deux fois moins. Ces différences s'expliquent en partie par la ségrégation résidentielle : si tous les collèges de chaque commune avaient exactement la même composition sociale, l'indice de ségrégation baisserait que de 25%. Le résidu est le résultat de deux phénomènes, que les données ne permettent pas de distinguer : la ségrégation résidentielle entre les quartiers de chaque commune, et la ségrégation liée directement aux établissements (par exemple par le biais des dérogations).

La ségrégation en fonction du niveau scolaire, elle, est beaucoup plus variable en fonction du niveau. Elle reste limitée au collège, où elle varie entre 7 et 9%. La part des « bons élèves » y est de 21%, mais un bon élève comptera en moyenne 27% bons élèves parmi les élèves du même niveau (sixième, cinquième, etc.) et du même collège, alors que les autres n'en comptent que 19%. Au lycée, la ségrégation scolaire augmente nettement, pour atteindre 18 à 21%. Cette

forte augmentation s'explique notamment par le début de la filiarisation : les élèves fréquentent des lycées différents selon qu'ils choisissent la voie générale, technologique ou professionnelle – à l'exception des 20% de lycées polyvalents.

Les indicateurs de ségrégation sont des valeurs moyennes qui cachent une grande variété de situations. On trouve également un petit nombre d'établissements qui accueillent un public soit très favorisé, soit très défavorisé.. Ainsi, en troisième, 10% des élèves fréquentent des établissements contenant 5% ou moins d'élèves CSP+ dans leur niveau ; à l'inverse, 5% des élèves ont plus de 60% d'élèves CSP+ dans leur cohorte, et même plus de 80% de CSP+ pour les 1% d'élèves (soit plus de 7 000 élèves) dont les environnements sont les plus favorisés. Si on s'intéresse aux élèves issus des milieux les plus populaires (ouvriers, chômeurs et inactifs), qui représentent 37% des élèves de troisième, 10% des élèves en comptent 63% ou plus dans leur établissement ; 5% en comptent 71% ou plus. De tels écarts sont également observés en termes de ségrégation scolaire : 10% d'élèves comptent 6% ou moins de « bons élèves » dans leur établissement, et à l'inverse 5% d'élèves en comptent plus de 43%, et 1% en comptent plus de 58%.

La ségrégation entre établissements a une amplitude très variable d'un département à un autre, tant dans la dimension sociale que scolaire (ces deux dimensions étant très corrélées). Par exemple, la ségrégation sociale varie de 2% à 27% ; les départements ayant la plus forte ségrégation sociale sont essentiellement des départements urbains qui comportent des grandes villes. Dans les départements à faible densité de population, les collèges recrutent sur un rayon pouvant dépasser les dix kilomètres : ils regroupent donc dans un même lieu des élèves d'origines différentes, ce qui favorise la mixité sociale. La multiplication du nombre de collèges, dans les zones urbaines, augmente au contraire la ségrégation par deux biais : d'abord, parce que les collèges reflètent plus précisément la ségrégation résidentielle, et ensuite parce qu'il s'installe une situation de concurrence qui fait émerger des collèges « souhaités » et des collèges « évités ».

Nous nous intéressons également dans cette étude au rôle de la composition des classes dans la ségrégation sociale et scolaire. Cette dimension, souvent absente dans les études quantitatives (par manque d'accès aux données et en raison des difficultés méthodologiques qu'elle pose), a une importance cruciale dans l'étude de la ségrégation. En effet, une politique volontariste de

déségrégation entre les établissements peut être compensée par une reségrégation à l'intérieur de ceux-ci. La ségrégation totale, qui est la somme de la ségrégation entre les établissements et de la ségrégation entre les classes de chaque établissement, risque alors de rester inchangée, et ce simple changement de structure de la ségrégation peut même avoir des effets néfastes. La ségrégation sociale entre les classes, au sein des établissements, varie entre 4 et 6 points selon le niveau. Cela signifie que la ségrégation totale entre l'ensemble des classes est supérieure à la ségrégation entre les établissements de 4 à 6 points. En classe de troisième par exemple, la ségrégation sociale entre établissements est de 17%, et la ségrégation entre les classes des établissements est de 5 points, soit un total de 22%. Ainsi, un élève CSP+ va compter une part d'élèves CSP+ *dans son établissement* supérieure de 17 points à celle connue par un élève non-CSP+, et une part d'élèves CSP+ *dans sa classe* supérieure de 22 points à celle connue par un élève non CSP+, soit 5 élèves sur une classe de 25. Globalement, la composition des classes a un effet relativement limité sur la ségrégation sociale, puisqu'elle ne représente que 20% de la ségrégation totale. Cependant, elle joue un rôle beaucoup plus important sur la ségrégation scolaire : *la ségrégation scolaire totale vaut le double de la ségrégation scolaire entre établissements*, c'est-à-dire que les compositions de classe contribuent autant à la ségrégation scolaire que la ségrégation résidentielle et la ségrégation entre établissements. Au total, l'indice de ségrégation scolaire totale vaut 13 à 18% au collège, 28% en seconde, 38% en première et 36% en terminale.

Alors que la ségrégation entre établissements est un phénomène plus prononcé dans les zones urbaines, la ségrégation entre les classes des établissements varie très peu d'un département à un autre, à l'exception notable des départements d'outre-mer où elle prend des valeurs beaucoup plus importantes qu'en métropole. La ségrégation entre les classes des établissements concerne à la fois les zones urbaines et les zones rurales, et vient s'ajouter à la ségrégation entre établissements, quelle que soit son niveau : il ne s'agit donc pas d'un mécanisme de compensation de la mixité mais d'un phénomène qui a lieu aussi bien dans des zones de mixité que dans des zones à forte ségrégation entre établissements.

Pour évaluer le niveau de ségrégation entre les classes des établissements, nous avons comparé les valeurs réelles des indices de ségrégation à des valeurs obtenues sur une série de 100 affectations aléatoires des élèves aux classes. Le hasard explique une partie non négligeable de la

ségrégation, en raison de la petite taille des classes. Mais en troisième, 25% des établissements ont une ségrégation sociale entre les classes qui dépasse ce qui s'explique par le hasard. Dans 45% des collèges, c'est la ségrégation scolaire qui dépasse cette ségrégation « aléatoire » obtenue par les simulations : près de la moitié des établissements créent des classes de niveau. Un instrument important de cette ségrégation au sein des établissements est le jeu des options. Les classes bilingues en sixième et cinquième, le latin à partir de la cinquième et les sections européennes à partir de la quatrième sont des options fortement marquées socialement et scolairement et sont des facteurs de différenciation des établissements et des classes. Les parcours bilingues et les sections européennes ne sont proposés que dans une moitié des collèges environ, à l'inverse du latin qui l'est dans presque tous les établissements. Dans les trois cas, les élèves choisissant ces options sont regroupés dans la moitié des classes en moyenne ; dans une minorité de cas, ils sont regroupés dans une ou deux classes seulement, ou répartis dans toutes les classes de l'établissement. Cela donne lieu, dans les établissements où ces options sont proposées, à une hiérarchie sociale et scolaire entre les classes.

Enfin, nous avons étudié l'évolution des indices de ségrégation sur une décennie au niveau national. Nous observons que la ségrégation sociale a très peu évolué pour les cohortes d'élèves ayant passé le DNB entre 2006 et 2014. Cette stabilité dans les chiffres au niveau national masque des progressions diverses entre les territoires ; plusieurs études ont notamment relevé une hausse de la ségrégation dans certaines agglomérations suite à l'assouplissement de la carte scolaire. Dans la même période, on constate une diminution sensible de la ségrégation scolaire au niveau national, à la fois entre les établissements et entre les classes au sein des établissements.

La mesure de la ségrégation dans la durée

Enfin, le chapitre 4 propose une méthode pour mesurer la persistance de la ségrégation dans la durée. Au cours de sa scolarité, un élève est susceptible de connaître des environnements différents : par exemple, il peut être dans une « bonne classe » en sixième et cinquième et dans une « mauvaise classe » en quatrième et troisième. L'existence de ces « bonnes » et « mauvaises » classes est la manifestation concrète d'une ségrégation qui peut exister à chaque instant. Cette ségrégation peut cependant être atténuée si les élèves qui fréquentent les « mauvaises classes » ne

sont pas toujours les mêmes. Dans ce chapitre, nous proposons des outils statistiques permettant de mesurer cette diversité des environnements fréquentés, c'est-à-dire la mobilité des élèves entre les environnements sociaux.

Des outils statistiques similaires existent déjà dans le cadre de la mesure des inégalités de revenus. Pour mesurer, à un instant donné, les inégalités de revenus, les statisticiens ont recours à un ensemble d'indices, le plus connu d'entre eux étant l'indice de Gini. Cet indice mesure l'écart moyen de revenus entre deux personnes choisies au hasard, divisé par le revenu moyen de la population considéré. Il varie entre 0 (absence d'inégalité) et 100% (inégalité totale : une seule personne reçoit tous les revenus). L'existence d'inégalités de revenus peut être atténuée, voire même totalement annulée par la *mobilité* des individus dans la distribution des revenus. Considérons une situation fictive où chaque individu a le même salaire en fonction de son âge (par exemple, un salaire croissant lié à l'expérience) et a la même espérance de vie. À chaque instant, on observerait une inégalité de revenus qui reflèterait tout simplement la pyramide des âges. Mais en considérant les revenus moyens gagnés par les individus au cours de leur vie, ceux-ci seraient identiques pour tout le monde : l'indice de Gini pour ces revenus moyens serait égal à zéro. C'est une situation de mobilité parfaite : des inégalités existent à chaque instant, mais celles-ci disparaissent lorsqu'on considère les revenus moyens sur une longue période. À partir de cette situation, on peut construire un indice de mobilité de revenus, qui est défini par le taux de diminution de l'indice d'inégalité lorsqu'on passe des revenus instantanés aux revenus moyennés dans le temps. Dans l'exemple précédent, puisqu'on part d'un niveau d'égalité non nul en utilisant les revenus instantanés et qu'on arrive à un niveau zéro d'inégalités en utilisant les revenus moyens, la diminution est de 100% : la mobilité est parfaite. En réalité, la mobilité de revenus est beaucoup plus faible. Aux États-Unis par exemple, elle est de l'ordre de 40% : l'indice de Gini calculé à partir des revenus moyens est 40% plus faible que l'indice de Gini calculé à partir des revenus instantanés.

Notre démarche consiste à adapter ce raisonnement au cas de la ségrégation. De la même manière que l'indice de Gini mesure les inégalités de revenus, plusieurs indices permettent de mesurer les inégalités de composition sociale de l'environnement, c'est-à-dire la ségrégation. Ces indices de ségrégation sont calculés à partir de la proportion d'un groupe donné (par exemple les

élèves issus de milieux aisés ou les élèves ayant obtenu les meilleures notes) dans l'entourage de chaque élève à *un instant donné*. Il est cependant possible de modifier ces indices pour utiliser non pas les environnements à un instant donné, mais la moyenne des environnements fréquentés pendant toute la scolarité au collège, par exemple. En mesurant la différence entre les deux indices de ségrégation ainsi obtenus, on obtient un indice de *mobilité entre environnements sociaux*. Cet indice peut également être interprété comme le degré de diversité des environnements sociaux fréquentés par un élève au cours de sa scolarité : s'il vaut zéro, chaque élève a toujours connu le même environnement ; plus il augmente, plus cela signifie que les élèves ont fréquenté des environnements différents.

L'indice de mobilité entre environnements sociaux a cependant une différence fondamentale avec l'indice de mobilité de revenus : il ne peut jamais atteindre 100% dès lors qu'il existe de la ségrégation instantanée. Ainsi, si un principal de collège crée des classes de niveau, il ne pourra jamais effacer totalement les inégalités en alternant d'une année sur l'autre les élèves qui y sont affectés. L'idée est la suivante : par définition, chaque année, il y a plus de places pour les meilleurs élèves dans les meilleures classes et plus de places pour les élèves plus faibles dans les classes plus faibles. Il y a donc une marge de manœuvre limitée et les élèves les plus faibles seront, en moyenne, plus souvent affectés à des classes faibles que leurs camarades de niveau plus élevé.

Notre article permet de préciser cette limite quantitativement : plus la ségrégation moyenne à chaque instant est élevée, moins la mobilité entre environnements sociaux pourra être élevée. La relation mathématique dépend de l'indice de ségrégation utilisé ; en considérant l'indice d'exposition que nous utilisons notamment dans le chapitre 3, celle-ci s'exprime particulièrement simplement : par exemple, si la ségrégation à chaque instant vaut en moyenne 30%, la mobilité ne pourra en aucun cas excéder 70% ($100 - 30$). Dans ce scénario de mobilité optimale, la ségrégation mesurée en utilisant les environnements moyens vaudrait alors 9%.

L'utilisation de cet indice de mobilité entre environnements sociaux peut poser quelques difficultés dans le cas où de l'attrition est observée, c'est-à-dire si certains élèves ne sont pas présents dans l'échantillon pendant les quatre années ; notre article propose une méthodologie pour calculer l'indice de mobilité entre environnements sociaux sur des données réelles, notamment en

présence d'attrition.

À titre d'illustration, nous avons calculé l'indice de mobilité entre environnements sociaux en l'appliquant à la ségrégation sociale au collège. L'indice de ségrégation sociale « instantané » est d'environ 20% dans les collèges français, comme nous le voyons en détail dans le chapitre 3. L'indice de mobilité entre environnements sociaux a une valeur proche de 15% : cela signifie que si on mesurait la ségrégation sociale en utilisant les environnements moyens fréquentés au cours des quatre années de collège, l'indice vaudrait 17% au lieu de 20%. Cette valeur de 15%, qui peut paraître faible, doit être lue avec précaution. Elle signifie qu'un collégien moyen connaît au cours de sa scolarité environ 15% de la diversité des environnements de classe existant *dans l'ensemble du pays*. Or, même si la mobilité pourrait, mathématiquement, être beaucoup plus élevée, elle impliquerait des déplacements d'élèves très importants, non seulement entre classes d'un même collège mais aussi entre collèges, entre villes, voire entre régions : en effet, les différences d'environnements sociaux entre établissements sont pour une part importante les conséquences d'inégalités territoriales qui ne sont pas le seul fait de l'institution scolaire. Notre outil permet cependant de mettre en place un pilotage à un niveau plus local, en calculant des indices de mobilité par région, par commune, voire même par établissement.

Chapitre 1

Elite Tutors and Underprivileged High-School Students

Ce chapitre a été écrit avec Son Thierry LY et Éric MAURIN.

Nous remercions Laurent DAVEZIES et Pascaline DUPAS pour leurs nombreux commentaires sur une version antérieure de cet article.

Cette recherche a bénéficié d'un soutien financier du fonds d'expérimentation pour la jeunesse.

1.1 Introduction

In most developed countries, students coming from low-income families are massively under-represented in the most prestigious programs of higher education. It contributes to the exclusion of entire social groups from political and economic elites.

This issue has attracted considerable attention from both policy makers and social scientists, but the mechanisms driving unequal access to higher education are still not well understood (Bailey and Dynarski, 2011). One potential explanation is that low-income high school students lack basic information about the higher education system and its selectivity. It may lead them to develop inadequate application strategies (Smith et al., 2013; Hoxby and Avery, 2013; Hoxby and Turner, 2013b).

In recent years, many initiatives have flourished around the world to address this issue. In France, the ministry of Education has encouraged institutions of higher education to develop joint programs with underprivileged high schools in order to help the best students from these high schools to form more ambitious plans and meet more ambitious goals. There exist about 350 such programs all over the country. They typically provide free tutoring to help high school students improve their academic achievement. They also provide students with assistance with college choices and applications. These programs are called "*cordées de la réussite*" (team for success) and are becoming increasingly popular. To the best of our knowledge, however, very little is known about the actual impact of these programs on eligible students. On the one hand, they contribute to bridge the informational gap between underprivileged students and higher education. But on the other hand, they are often time consuming and may contribute to distract students from basic subjects.

To shed light on this issue, our paper develops a randomized evaluation of the impact of one of the oldest "*cordée de la réussite*". It is operated by the most prestigious and selective institution of higher education in France, namely the *École normale supérieure* (hereafter, ENS). The ENS encompasses both a very selective graduate school and a set of world class research centres. In Math, the ENS is for example the second institution in the world (just below Princeton) in terms of Field medals won by former students. This institution has played a leading role in the selection and training of French intellectual elite for more than a century.

Since 2006, the ENS offers each year a two-year mentoring and tutoring program (about 150 hours per student and year) to a selection of promising students coming from twelve underprivileged high schools of the Paris Region. The tutoring is provided by graduate students

from the ENS itself. The set of eligible students is defined at the end of the first year of high school (grade 10), while the program itself lasts for the two last high school years (grade 11 and 12). Each year, in each school, eligible students are randomly selected from a set of volunteer students identified by school principals as having the potential to go to higher education. After the randomization, the set of eligible students is divided in groups of six students and each group is randomly assigned to a specific ENS tutor. In grades 11 and 12, these groups of eligible students are invited each month to participate in one or two tutoring sessions dedicated to either exploring new fields of study, deepening subjects or preparing high school exit examination. In grade 12, eligible students are also invited to participate in specific sessions dedicated to help them apply to higher education institutions. The travelling costs involved by weekly participation in the program (as well as the cost of attending a one-week introductory meeting at ENS in Paris) are all covered by the program. All in all, the intervention costs about 1,500 euros per student and year.

This paper focuses on volunteer students identified by principals in 2010 and 2011. It shows the results of comparing the achievement and choices of those randomly selected to be eligible (the treatment group) with the achievement and choices of those not selected (the control group). This evaluation reveals that the intervention has very little effects on students' average outcomes. In particular, there is no significant difference between treatment and control groups in average performance on the national high school exams (*baccalauréat*) taken at the end of grade 11 and grade 12. Similarly, there is no significant difference in the proportion of students who get access to (and are able to persist in) the most selective undergraduate programs (called *Classes Préparatoires aux Grandes Écoles*, hereafter CPGE).

However, further explorations reveal that the effect of the intervention is very different across ability groups. Among the 50% eligible students with the highest level of achievement pre-treatment, the intervention induces a significant increase in high school achievement as well as in the probability to get access to (and persist in) the most selective undergraduate programs. By contrast, among the 50% eligible students with the lowest level of ability pre-treatment, the intervention induces a significant decrease in both high school achievement and probability to enter into selective undergraduate programs. All in all, the intervention has no impact on the

average outcomes of eligible high school students, but a very important effect on inequalities across eligible students with different ability levels pre-treatment. The gap in high school graduation between ability groups is three times more important in the treatment group than in the control group.

Our paper contributes to the recent and growing literature that explores the impact of providing high school students from underprivileged background with assistance to gain access to higher education (see *e.g.* Avery, 2009, 2013; Bettinger et al., 2012; Castleman and Page, 2015; Hoxby and Turner, 2013a). One distinctive feature of the intervention analyzed in this paper is that it does not simply try to improve students' information about higher education (or about how to apply to higher education institutions), but also their academic readiness for higher education, which explains the large number of sessions and their academic content. Another important feature of our experiment is that it is not about the US system, but about a European system, where higher education is almost free and where tuition fees are not an issue. In this context, our most striking result is maybe that such an intervention can have negative effects on a significant fraction of students, even when targeted on those who volunteer and are identified as having the potential to go to higher education. Many high school students with underprivileged background find it pleasurable to spend a lot of time participating in a program which, in the end, happens to have a negative effect on their performance on key exams and a negative effect on their probability to gain access to (and persist) in the best undergraduate programs. In 2016, we conducted a post-treatment satisfaction survey on the sample of students who participated in the program in 2010 and 2011. It confirms that the vast majority of former participants found it enjoyable to participate in the program both because of the quality of their tutor and the quality of their group of tutees. But it also confirms that many participants perceived the participation in the program as very time consuming and difficult to reconcile with school homeworks. The overall outcome of participation in the program turned out to be very positive for the higher-ability participants, but very negative for the lower-ability ones.

The paper is organized as follows. The next section describes the context of the experiment, the content of the intervention and the randomization process. Section 1.3 describes the data used in the econometric analysis while section 1.4 presents the main results of our experiment,

namely the impact of the intervention on students' performance on high school exit exams as well as on their probability to gain access to (and persist in) selective undergraduate program. Section 1.5 builds on the fact that tutors were randomly assigned to groups of tutees to explore the variation in the effect of the program across the different types of tutors. Section 1.6 builds on the survey on former participants conducted in 2016 to test assumptions about why a program designed to help students may end up having a negative effect on a significant fraction of them.

1.2 The experiment

1.2.1 Institutional context

In France, compulsory education encompasses 5 years of elementary education (between age 6 and 11) and 4 years of middle school (between age 11 and 15). At the end of middle school (grade 9), about 60% students pursue general education in high school whereas 40% go to a vocational school or enter into the labour market. One year later, at the end of grade 10, about 70% of the students who are still in the general education track (*i.e.*, about 40% of a birth cohort) pursue general education in high school for two additional years (grades 11 and 12) in order to prepare a general education diploma. The other 30% enter into a more technical education program. Students who pursue general education at the end of grade 10 have on average a much better academic level than those who pursue technical or vocational education. At the end of grade 10, students pursuing general education have to choose a major field of study and about half of them specialize in science whereas the other half specialize in humanities (either in Literature/Languages or Economics/Social Sciences). As discussed below, students eligible to the ENS tutoring program are selected among the high-achieving 10th graders of 12 underprivileged high schools. Virtually all of them pursue general education at the end of grade 10 and about two thirds specialize in science.

In grades 11 and 12, high school students prepare for the different exams required for high school graduation. There is one exam per subject and graduation is based on the average mark across the different subjects (the average has to be above 10/20). Graduation is a necessary condition for entering into higher education. Some specific exams take place at the end of grade

11 (most notably oral and written French exams), but most of them take place at the end of grade 12. The overall number of exams and the relative importance of the different subjects depend on whether the student chose to specialize in science or in humanities at the end of grade 10.

Students who want to enter into a selective undergraduate program have to apply through a Centralized Assignment System (called *Admissions Post-Bac*, hereafter APB). They are asked to list up to 36 specific programs by descending order of preference. Each program then ranks its applicants based on the marks obtained during 11th and 12th grade (as well as on high school teachers' assessment). Using a deferred acceptance mechanism (Roth, 2008), the system then assigns as many students as possible to one of their listed choices.

The most selective undergraduate programs correspond to the *Classes Préparatoires aux Grandes Écoles* (CPGE). Among students who prepare for a general education high school diploma, about 13% are admitted in a CPGE (17% of those who specialize in science), and 77% make it to the second year. These programs involve two years of intense preparation (either in science or humanities) at the end of which students take competitive exams for entry into the most prestigious graduate programs. Admission to the ENS itself is based on one such competitive exam. Most ENS students have gone through a two-year CPGE preparation program before entry into ENS. About 22% of second-year CPGE students repeated that year in order to take the competitive exams again.

1.2.2 The program and its objectives

In 2008, the French government initiated programs (called "*cordées de la réussite*") all over the country in order to increase the proportion of students from underprivileged high schools entering higher education programs. At the local level, each specific program corresponds to the association of an institution of higher education with a set of high schools located in the same region. Each year, a selection of students from these high schools are provided with information on higher education institutions (selectivity, fields of study, etc.), counsels on how to apply to these institutions and assistance to improve their high school academic achievement. In most cases, this assistance is provided by volunteer tutors coming from the higher education institution itself.

The program analyzed in this paper corresponds to the network constituted by the *École normale supérieure* (ENS) and twelve underprivileged high schools from Paris and its region.¹ The high schools were selected based on the socioeconomic background of their students as well as on the proportion of students they send to CPGE programs: only 8% of their students enter into such selective programs (11% of those who specialize in science) which is about two times less than the average high school in the region of Paris (14%, and 20% of those who specialize in science). As discussed above, the ENS represents one of the most prestigious institutions of graduate education in France. It encompasses several highly selective graduate schools as well as top research centers.

1.2.3 Identification of volunteers

Each year, in each school, participants to the program are selected from a set of 10th grade students identified as volunteers. The identification of volunteers takes place early April, about two month before the end of the academic year. In each school, the principals starts by identifying 10th grade students who are very likely to pursue in the general education track and invite them to participate in an information meeting where the managers of the ENS program provide them with information on the objectives and contents of the program. At the end of this information meeting, students who are interested in actual participation are invited to take a questionnaire (about their family background and school experience), to fill it at home and to bring it back one week later. Those who come back one week later with their filled questionnaire have a short interview with the program manager. The manager checks whether the questionnaire is well filled and also whether the student has well understood the implications of participation. In particular, students are reminded that not all volunteers will eventually be eligible to participate in the program, only a random selection. At the end of this interview, students who confirm their willingness to participate in the program are considered as volunteers.

On a typical year, the capacity of the program is about 140 seats and there are about 200 volunteers. On academic year 2009-2010 (our first cohort), the ENS agreed to temporarily

1. The Parisian region is divided into three education districts (these districts are called *académies*), namely the district of the city of Paris itself (*académie de Paris*), the eastern district (*académie de Créteil*), and the western district (*académie de Versailles*). We have five schools in the district of Paris, six schools in the eastern district and one school in the western district.

TABLE I.I – The characteristics of students who volunteer to participate in the program

	Paris region	Experimental high schools	Volunteer students	Volunteers, cohort 2010	Volunteers, cohort 2011
Variable	(I)	(II)	(III)	(IV)	(V)
Female	0.528	0.536	0.600	0.613	0.567
High-SES	0.416	0.200	0.253	0.257	0.242
Low-SES	0.188	0.344	0.274	0.262	0.306
Ability score (grade 9)	0.000	−0.556	0.153	0.132	0.208
Repeat grade 10	0.118	0.131	0.000	0.000	0.000
Pursue general education	0.583	0.533	0.996	1.000	0.987
Specialize in science	0.304	0.236	0.634	0.586	0.758
Obs.	209,654	7,025	558	401	157

Note: Column (I) shows the average characteristics of general education 10th grade students in the region of Paris. Column (II) shows the same characteristics for general education 10th grade students in the twelve high schools of the experiment and column (III) for the volunteer students in these twelve high schools. Columns (IV) and (V) further show the characteristics of volunteer students in each cohort. The ability score corresponds to the (standardized at the Paris region level) average grade obtained at the national middle school exit exams taken at the end of grade 9.

Reading: 60% of volunteer students are female, 25.3% come from a high SES family background, 0% are repeating grade 10. At the end of grade 10, 99.6% choose to pursue general education, 63.4% choose to specialize in science. Their average standardized score at the end of middle school exams (*i.e.*, grade 9) is 0.786.

increase the capacity of the program and school principals were encouraged to boost participation in the information meetings. On this specific year, we ended up with 401 volunteers for 221 seats. On the following year (2010-2011), the capacity was back to normal (137 seats) and the number of volunteers back to 210 students.

1.2.4 Random selection of eligible students

Once the lists of volunteers were completed in the different experimental schools, eligible students were randomly selected from these lists. As discussed below, the randomization was stratified by school and main field of study. In 2011, 53 volunteer students did not participate in the random draw (and were automatically selected) because there were too few volunteer students

in their school with similar field of study. Overall, for the two cohorts under consideration, a total of 558 volunteer students were included in the random draw of eligible students, 401 on year 2010 and 157 on year 2011.

Building on administrative data that are described below, **TABLE 1.1** provides some basic statistics on these volunteer students as well as on their non-volunteer schoolmates and on the average students in the Parisian region. The Table confirms that volunteer students have a much better academic record not only than their non-volunteer schoolmates, but also than the average high school student in the Parisian region, even though they come more often from a low-SES family than the average high school student. Specifically, the score obtained by volunteer students at the end-of-middle-school national examination is about +15% of a SD higher than the score obtained by the average student in the Parisian region, even though the proportion of students coming from a low-SES family is about 50% more important among volunteer students than among the average student in the Parisian region. Virtually all volunteer students pursue in the general education track in grade 11 whereas the average proportion in the Parisian region is only about 63%. Generally speaking the Table confirms that the program was able to target relatively high-ability students with relatively low socioeconomic background, compared to the average student in the Parisian region.

For both cohorts, the randomization took place in each school just before the start of the summer holidays and just after the principal pre-assigned each volunteer student to one future 11th grade class (based on her choice of field of study). In French high schools, each 11th grade class corresponds to either students who specialize in science at the end of grade 10 or to students who specialize in humanities (*i.e.*, social sciences or languages/literature). The randomization was conducted at the class level and stratified by major field of study (science/humanities). For each major field of study, half of the classes – or half rounded up to the nearest integer when there were an odd number of classes – were put in the treatment group. In the end, we have 307 volunteer students in the treatment group and 251 volunteer students in the control group.

Only the 307 volunteer students in the treatment group were eventually invited to participate in the program. Most of the results in this paper are based on the comparison of volunteer

students in treatment and control groups. Under the assumption that volunteer students in control groups remain unaffected by the treatment (SUTVA), this comparison provides an estimate of an intention-to-treat parameter, namely the impact of being invited to participate in the program on the subsequent outcomes of volunteer students.

To assess the similarity between the control and treatment groups, Appendix [TABLE 1.A.1](#) compares the responses of treatment and control groups to the questionnaire that students had to fill in order to be identified as volunteers.² Comfortingly, we find little difference in responses across the treatment and control groups. Specifically, differences between treatment and control groups are statistically different at the 10% level for one survey item only out of 22 (and significant at the 5% level for no item).

To further test for the similarity of the two groups, Appendix [TABLE 1.A.2](#) (panel A) builds on the information provided by administrative registers to compare the socio-demographic characteristics of volunteer students in the treatment and control groups (in terms of gender, grade repetition, parental occupation, pre-treatment grades). Again, we find no significant differences between the two groups.

To test for heterogeneous treatment effects, most of our regression analysis will be conducted not only on the full sample, but also separately on the subsample with the highest academic results pre-treatment (*i.e.*, in 10th grade) and on the subsample with the lowest academic results. We replicate the comparison of the pre-treatment characteristics of treated and control students separately for the two subsamples in panels B and C of [TABLE 1.A.2](#). We do not detect any significant pre-treatment difference across treatment and control students within both ability groups.

2. This survey provides information on students' family background, main topics of interest, school background, plans for the future, level of information about higher education institutions, type of help they can get for making choice. This information is available for 93% of the students identified as volunteer in either 2010 or 2011. The missing rate is very similar in control and treatment groups.

1.2.5 Program content, tutors and take up

Generally speaking, the objective of the ENS program is to provide volunteer students from underprivileged high school with information on post-secondary education programs (especially the most selective ones) as well as to improve their readiness for these programs. Just after the randomization, the program managers sent a letter to the students who participated in the random draw in order to inform them of the results of the draw and to invite those of the treatment group to participate in a kick-off week (called the campus week) organized at ENS late August, just before the start of the following academic year.

During this campus week, students have activities led by tutors from previous cohorts. The objective of these activities is to help students improve their methods of work and to prepare them for grade 11. Team building activities are also organized in the evening. Meals and accommodation are paid for by the program. In 2010 and 2011, the vast majority of students did participate to this kick-off week (see [TABLE 1.2](#)).

After the campus week, students are asked to choose a theme for the tutoring sessions in which they are going to be involved throughout the academic year. There are four possible themes: science, social sciences and history or literature. Students are also asked to list one or two friends with whom they would like to be grouped for these sessions. These lists are used by program managers to define triplets of students with similar thematic preferences. The managers then randomly paired triplets within each theme and randomly assigned each pair of triplets to one of the tutors specialized in its theme.

Each year, about 80% of tutors are new ones whereas 20% have already been involved in the program. New tutors are recruited early October at the beginning of the academic year. They are all ENS students. They first benefit from a two-day training session where program managers provide them with information on the objectives of the program and on the type of high school students they are going to mentor. Tutors who have already been involved in the program also participate in this training session in order to share their experience. The tutoring program starts at the end of October, with a first meeting between tutors and their groups of six students at ENS. The tutors meet three times a year in order to share experience and get feedback from

TABLE I.2 – Take up rates, by cohort and ability group

	All	Lower ability	Higher ability
	(I)	(II)	(III)
Panel A: Cohort 2010			
Introductory week (year 1)	0.786	0.694	0.881
More than 10 sessions attended (year 1)	0.639	0.615	0.667
Re-enlistment	0.559	0.459	0.661
Introductory week (year 2)	0.282	0.252	0.312
Obs.	220	111	109
Panel B: Cohort 2011			
Introductory week (year 1)	0.724	0.721	0.727
Re-enlistment	0.483	0.372	0.591
Obs.	87	43	44

The sample includes eligible students from cohorts 2010 (panel A) and from cohort 2011 (panel B). In each panel, column (I) shows take up rates for the full sample whereas column (II) (column (III)) shows the take up rates for the subsamples of students with lower (higher) pre-treatment ability score.

Reading: Among eligible students from cohort 2010, 78.6% attended the (year 1) introductory week at ENS, 63.9% attended more than 10 tutoring sessions during the first year of the program, 55.9% re-enlisted at the end of the first year and 28.2% attended the (year 2) introductory week at ENS.

For cohort 2011, information on attendance was not collected, and there was no introductory week in year 2.

program managers.

The academic year is divided in three terms (September-December, January-March and April-June). During the first year of treatment, students from cohort 2010 benefited from four thematic tutoring sessions per term. An additional session of personal coaching was organized in the first term, as well as a session of improvisation theatre in the second term. All in all, students from cohort 2010 benefited from 12 tutoring sessions and two additional activities during the first year. This number of sessions was deemed excessive by a number of students, and it was reduced for the second cohort to two tutoring sessions per term, with only one additional cultural outing between the two sessions of each term, that is a total of 9 sessions for the second cohort (instead of 12 for the first cohort). We have information (collected by tutors) on attendance at the 12 tutoring sessions organized for the first cohort of students. This information is suggestive that attendance rates were high: about two third of students participate in 10 (or more) sessions out of 12 ([TABLE 1.2](#)).

At the end of grade 11, students are asked whether they want to pursue the program in grade 12. About 54 % of them chose to do so. Students from the first cohort were invited to participate to a second campus meeting. This second campus meeting is organized at ENS jointly with the first campus of the next cohort. However, students are not invited to stay at ENS during the campus (only lunches are covered). 50 % of the students who reenlisted came to the second campus.

At the start of the second year of treatment (grade 12), students can change tutoring groups and theme. As far as students from the first cohort (2010) are concerned, they benefited during this second year from seven thematic sessions with their tutors, as well as from one cultural outing and from a one-day forum of information on higher education. Also, starting in March, six additional sessions were organized in order to help students prepare the *Baccalaurate* exams taken at the end of grade 12. As far as students from the second cohort are concerned, they benefited during the second year of treatment from four thematic sessions (and two cultural outings) during the two first terms of the year. They also benefited from two additional sessions during the third and last term of the year, dedicated to the preparation of the *Baccalaurate* exams.

1.3 Data

The high schools participating in the experiment provided us with the ID number of their volunteer students as well as their average academic achievement in grade 10 (which we will use to define our two basic ability groups). Using ID numbers, we first augment this dataset using exhaustive administrative data on students' performance on the national exam taken at the end of 9th grade (exam called the *Diplôme National du Brevet*, hereafter the DNB) as well as on the national exam taken at the end of 12th grade (the *Baccalauréat*). We were also able to augment our initial dataset using administrative school registers (*Bases Centrales Scolarité*, hereafter the BCS) which provided us with information on the field of study chosen by students at the end of grade 10 as well as on the schools and classes that they attend in the following years.

Overall our working sample consists of the 558 tenth-graders who volunteer to participate in the program in either 2010 or 2011 with information on their treatment status, their ability level in grade 10, their pre-treatment performance at the middle school national exam (in grade 9) as well as information on their future performance at the high school exit exams (in grades 11 and 12). Generally speaking, missing rates are very small and unrelated to students' treatment status. In particular, information on high school graduation (or on high school graduation with honors) is available for 99% of the observations (7 missing) whereas information on pre-treatment performance on grade 10 is available for 97% of the observations (16 missing).

Finally, to assess the similarity between the control and treatment groups, we were also able to use the information coming from the questionnaire that students had to fill in order to be identified as volunteer. This pre-treatment survey provides us with information on students' family background (parents' country of birth and date of arrival in France, parents' occupation and education, age and education of siblings, native language...), their preferred activities (sport, art...), their main topics of interest (politics, economics, literature, technology, etc.), their school record and school background (elective courses chosen, self-assessment, favorite subjects, whether they can get help for their homework), their plans for the future (field of study, institution of higher education, occupation), their level of information about higher education institutions, the type of help they can get for making choices. There is also a question about their opinion about what

equal opportunities should be. This information is available for 93% of the students identified as volunteer in either 2010 or 2011. Appendix [TABLE 1.A.1](#) presents descriptive statistics from this survey, as well as a balancing test showing that responses were very similar in the treatment and control groups. Differences between treatment and control groups are statistically different at the 10% level for one survey item only out of 22 (and significant at the 5% level for no item).

1.4 Effects on achievement and choices

In this section, we analyse the effect of the intervention on students' academic performance on high school national exams at the end of grade 11 and grade 12. We focus on exams (externally set and marked) that students have to take in order to obtain their high school national diploma (*baccalauréat*). For each exam, we estimate the following model:

$$Y_i = T_i \cdot \alpha + X_i \cdot \beta + \nu_i \quad (1.1)$$

where, for each student i , variable Y_i represents the mark obtained at the exam (or a dummy variable indicating whether i passed the exam), variable T_i is a dummy indicating whether i is in the treatment group, and X_i is a vector of pre-treatment control variables that includes dummies for gender, grade repetition, family background, pre-treatment marks as well as school fixed effects and major choice fixed effects. Variable ν_i represents unobserved error terms. The parameter of interest is α . Identification is a direct consequence of the experimental nature of the treatment assignment variable T . Standard errors are clustered at the class level.

1.4.1 Effects on high school achievement

[TABLE 1.3](#) shows the effect of the intervention on the grades obtained at the high school exams taken at the end of 11th grade (column I) as well as on the final grades obtained at the end of 12th grade (column II). It also shows the effect on the probability to graduate (column III) as well as on the probability to graduate with honors (column IV). The first panel refers to the full sample of volunteer students and we do not find any significant effect on the different outcomes in this sample. The intervention has no impact on the average grades obtained at the *baccalauréat* examinations by volunteer students nor on their high school graduation rate.

TABLE I.3 – The effect of the treatment on performance on high school exit examinations

Dependent variable	Grade 11 average score	Average score	Graduation	Honours
	(I)	(II)	(III)	(IV)
Panel A: All volunteer students				
Treatment	−0.081	0.012	0.020	0.012
s.e.	(0.078)	(0.080)	(0.041)	(0.045)
Obs.	537	537	551	551
Mean dep. var.	0.662	0.428	0.727	0.466
Panel B: Lower ability				
Treatment	−0.244**	−0.208	−0.077	−0.091*
s.e.	(0.114)	(0.131)	(0.065)	(0.053)
Obs.	264	264	273	273
Mean dep. var.	0.301	0.156	0.661	0.339
Panel C: Higher ability				
Treatment	0.059	0.240**	0.118**	0.108*
s.e.	(0.110)	(0.104)	(0.048)	(0.063)
Obs.	273	273	278	278
Mean dep. var.	1.019	0.698	0.792	0.592
Panel D: Differential impact				
<i>Treatment × higher ability</i>	<i>0.303*</i>	<i>0.448**</i>	<i>0.195**</i>	<i>0.199**</i>
s.e.	(0.159)	(0.176)	(0.078)	(0.084)
Obs.	537	537	551	551
Mean dep. var.	0.718	0.542	0.131	0.253

The sample includes volunteer students from cohorts 2010 and 2011.

The table shows the results from reduced-form regressions in which variables measuring performance on high school exit examinations (*baccalauréat*) are regressed on a treatment dummy, using students' gender, pre-treatment ability score and socioeconomic family background as control variables.

Column (I) shows the estimated effect of the treatment when the dependent variable is the average grade obtained at examinations taken at the end of grade 11 and column (II) when the dependent variable is the average grade across all examinations. Column (III) shows the estimated effect when the dependent variable is a dummy indicating high school graduation (average grade is 10/20 or more) and column (IV) when the dependent variable is a dummy indicating high school graduation with honours (average grade is 12/20 or more).

The first panel refers to the full sample, the second panel to subsample with lower pre-treatment ability score, the third panel to the subsample with higher pre-treatment ability score. Finally, the fourth panel shows the estimated difference in the effect of the treatment between lower and higher ability students. To estimate this difference, we use the full sample and regress the dependent variable on the interaction between a dummy indicating treatment and a dummy indicating high ability, controlling for the treatment dummy as well as for the full set of socio-demographic controls and their interactions with the higher ability dummy.

Standard errors clustered at the class level are reported in parentheses.

** and * denote significance at the 5% and 10% levels, respectively.

The second panel of [TABLE 1.3](#) refers to the subset of volunteer students whose pre-treatment academic achievement (in grade 10) is above the median whereas the third panel refers to the subsample whose pretreatment achievement is below the median. They reveal that the intervention has completely different effects on the two subpopulations: it contributes to a significant decrease in the grades obtained by the lower-ability group and to a significant increase in the grades obtained by the higher-ability one. The last panel confirms that the differences between the two sets of impacts are significant at standard level: the intervention contributes to a significant increase in the academic gap between the higher- and lower-ability groups. Specifically, differences in grades obtained at the end of grade 11 or grade 12 as well as differences in high school graduation probability between the two ability groups are significantly more important in the treatment group. For example, the difference in high school graduation probability between lower- and higher-ability students is about 13 percentage points in the control group, but this gap becomes about +20 percentage points larger in the treatment group, namely a tripling of the gap within the group that benefited from the intervention. Appendix [TABLE 1.A.3](#) shows the results of replicating this basic analysis separately on the two successive cohorts. It shows that results are qualitatively similar results for both cohorts, namely no effect of the treatment on average outcomes, but an increase in the gap between lower- and higher-ability students within the treatment group. For example, the difference in the probability of high school graduation between lower- and higher-ability students increases by about 19 percentage points in the first cohort and by about 25 percentage points in the second cohort. Because of the small size of the second cohort, it is not possible, however, to assess whether the increase in the gap in graduation rates between lower- and higher-ability students is more significant in the first or in the second cohort.

Higher-ability students chose more often to specialize in science and one explanation for our results may be that tutoring is more efficient for students who specialize in science. To test for this assumption, Appendix [TABLE 1.A.4](#) provides a replication of [TABLE 1.3](#) separately for the subsample of volunteer students who specialize in science and for those who specialize in humanities. The estimated gap in the effect of the intervention between lower- and higher-ability students appears to be even stronger for the group of students who specialize in humanities (although the variation in the gap across science and humanities is not significant at standard

level). Hence, the variation in the estimated impact of the intervention across lower- and higher-ability students does not seem to be driven by tutoring being more efficient in science.

In Appendix [TABLE 1.A.5](#) and [TABLE 1.A.6](#), we show the impact of the intervention on the grades obtained at the high school national examination in each subject (French, Math, Physics, Languages, etc.), major choice (Science/Humanities) and ability group. For students who specialize in science, the intervention contributes to a significant increase in the gap between lower- and higher-ability students in French written exam (+42% of a SD), Sciences (+65%), History (+42%) and Languages (+46%). The gap also increases in Math, Physics or Philosophy, even though these increases are not significant at standard level. Similar increases are also perceptible for students who specialize in humanities. Overall, the increased gap in achievement between lower- and higher-ability groups is not driven by a specific subject, but perceptible in virtually all subjects. These findings further confirm that the intervention did not affect subject-specific inputs, but more general determinants of performance at school. Overall, the simplest interpretation for the increased academic gap between lower- and higher-ability students is that participation in the program is time consuming and that only the more able students (pre-treatment) can afford spending a significant part of their time in activities that are not directly related to the preparation of the high school national examinations. Participation in the program involves additional time constraints that are detrimental in all subjects, especially for lower-ability students. Symmetrically, it likely contributes to an increase in motivation and educational aspirations, which is beneficial in all subjects, especially for higher-ability students.

1.4.2 Effects on access to selective undergraduate programs

One of the objectives of the program was to increase the proportion of students who gain access to the most selective undergraduate programs in France, namely the *Classes Préparatoires aux Grandes Écoles* (hereafter, CPGE program). It may be that the intervention has no average effect on high school grades, but contributes nonetheless to increase the number of students from underprivileged high schools who are aware of the existence of CPGE program and aspire to get admitted into one of them. According to our baseline survey, only about 49% of students

who volunteer to participate in the program know about CPGE. By contrast, most tutors get access to ENS after two years spent in a CPGE program and it is certainly the undergraduate program that they know the best.

To shed light on these issues, **TABLE 1.4** shows the effect of the intervention on the proportion of students who get access to (and attend year 1 of) a CPGE program after high school graduation (column I) as well as on the proportion who are still in a CPGE two years after high school graduation (column II). As the intervention has no effect on students' average high school grades, any effects on entry into CPGE would likely reflect an increased level of educational aspiration among eligible students. But we find no significant effect on the proportion attending year 1 of a CPGE program. This result likely reflects that the intervention has no effect on volunteer students' awareness of the existence of CPGE nor on their willingness to get admitted into these classes. Also we do not find any effect on the proportion attending year 2 of a CPGE program, consistent with the assumption that the program has no effect on the ability to persist in this type of program.

When we replicate this analysis separately on the lower-ability group, we find a significant negative effect on the proportion of students who attend year 1 (−6.8 percent points). This very strong negative impact likely reflects the negative effect of the program on high school achievement for this subgroup. We find a similar negative effect on year 2 as on year 1 attendance, consistent with the assumption that the intervention did not affect ability to persist for this lower-ability subgroup.

When we replicate this analysis on the higher-ability group, we find positive effects on both year 1 and year 2 attendance, but not significant at standard level. This slight improvement is likely driven by their increased academic performance in high school. Overall, there is no effect on the overall proportion of eligible students in CPGE, but a significant increase in variation in this proportion across lower- and higher-ability students (+12.7 percent points increase).

There are two basic types of CPGE, one specialized in science and one specialized in humanities. Columns III to VI explore whether the effect of the intervention is different across

TABLE I.4 – The effect of the treatment on access to selective undergraduate programs (CPGE)

<i>Bac</i> major	All majors		Science major		Humanities major	
	CPGE 1	CPGE 2	CPGE 1	CPGE 2	CPGE 1	CPGE 2
	(I)	(II)	(III)	(IV)	(V)	(VI)
Panel A: All volunteer students						
Treatment	−0.014	−0.015	0.011	−0.005	−0.062	−0.032
s.e.	(0.030)	(0.023)	(0.040)	(0.032)	(0.042)	(0.027)
Obs.	551	551	351	351	200	200
Mean dep. var.	0.165	0.112	0.156	0.123	0.179	0.095
Panel B: Lower ability						
Treatment	−0.065**	−0.057**	−0.065	−0.073**	−0.074*	−0.045*
s.e.	(0.030)	(0.022)	(0.041)	(0.034)	(0.043)	(0.026)
Obs.	273	273	131	131	142	142
Mean dep. var.	0.137	0.081	0.105	0.088	0.164	0.075
Panel C: Higher ability						
Treatment	0.056	0.037	0.073	0.046	−0.029	0.005
s.e.	(0.054)	(0.038)	(0.058)	(0.043)	(0.129)	(0.066)
Obs.	278	278	220	220	58	58
Mean dep. var.	0.192	0.144	0.186	0.144	0.214	0.143
Panel D: Differential impact						
<i>Treatment</i> × <i>higher ability</i>	0.121**	0.094**	0.138**	0.119**	0.045	0.050
s.e.	(0.060)	(0.044)	(0.069)	(0.053)	(0.116)	(0.063)
Obs.	551	551	351	351	200	200
Mean dep. var.	0.055	0.063	0.081	0.056	0.050	0.068

The sample includes volunteer students from cohorts 2010 and 2011.

The table shows the results from reduced-form regressions in which the probability to enter into a *Classe Préparatoire* program (CPGE1) or to gain access to the second year of a *Classe Préparatoire* program (CPGE2) is regressed on a treatment dummy, using students' gender, pre-treatment ability score and socioeconomic family background as control variables.

The first panel refer to the full sample whereas the second panel refers to subsample with lower pre-treatment ability score, the third panel to the subsample with higher pre-treatment ability score. Finally, the fourth panel shows the estimated difference in the effect of the treatment between lower and higher ability students. To estimate this difference, we use the full sample and regress the dependent variable on the interaction between a dummy indicating treatment and a dummy indicating high ability, controlling for the treatment dummy as well as for the full set of socio-demographic controls and their interactions with the higher ability dummy.

Within each panel, columns (III) and (IV) refer to students who specialize in science at the end of grade 10 whereas columns (V) and (VI) to students who specialize in humanities.

Standard errors clustered at the class level are reported in parentheses.

** and * denote significance at the 5% and 10% levels, respectively.

these two types of programs. It shows that the increased gap in CPGE attendance across ability groups is mainly driven by the scientific programs. For humanities, we observe a decline in attendance for both higher- and lower-ability students (although not significant at standard level for the higher-ability group) and no significant change in the gap.

1.5 Mechanisms: the role of tutors and peers

The previous sections are suggestive that the program has positive effects on volunteer students with the best academic level pre-treatment, but negative effects on the other volunteers. One important question, however, is whether the intervention induces the same positive and negative effects regardless of the tutors recruited to implement the program. Had all tutors the same negative effect on lower-ability students, it would be suggestive that this negative effect is related to some of the deep features of the program, its being too time consuming for instance. By contrast, would the negative effect be found for some specific tutors only, the problem would likely lie in the way the program is implemented rather than in its deep characteristics. A better selection (or training) of tutors would be a way to improve the program.

1.5.1 Tutors' characteristics

To explore this issue, it is possible to build on the fact that students were randomly assigned to their first-year (grade 11) tutor. In this set-up, the difference in outcomes observed at the end of grade 11 between volunteer students assigned to different types of tutors likely provides an evaluation of the effect of tutors. We have information on the gender of tutors, their family socio-economic background, the time elapsed since their admission at ENS and the design of our intervention makes it possible to test whether it makes a difference to be assigned to one type of tutors rather than to another one.

To implement this test, **TABLE 1.5** (panel A) focuses on volunteer students and show the results of regressing the average grade that they obtained at exams taken at the end of grade 11 on a set of dummy variables describing tutors' gender (male/female), their ENS admission year (ENS freshman/more senior ENS students) and their family socio-economic background (low

SES/high SES) controlling for the same basic set of pre-treatment variables as in the previous regression analysis. Regressions are conducted on the full sample of volunteer students from the first cohort (column (I)) as well as separately on the lower- and higher-ability groups (columns (II) and (III)), since the effects of the intervention are very different across the two groups.

The regression results do not show any significant effect of the gender of the tutor. By contrast, they reveal that the program induces significantly better outcomes when it is implemented by freshmen rather than by more senior tutors. This result holds true for both higher- and lower-ability participants. Within both ability groups, the difference in outcomes between recipients assigned to a freshman and recipients assigned to a more senior tutor is actually larger than the average difference in outcomes between treated and control students shown in [TABLE 1.3](#). These results suggest that the efficiency of tutors does not really increase with their level of experience in the program.³ The enthusiasm and motivation of beginners appear to be a key ingredient for the success of the program.

The regression results also show that recipients whose tutors come from a higher SES family background tend to obtain significantly better results than those whose tutors come from a lower SES family background. Again these results hold true for both higher- and lower-ability recipients, although the effect is statistically significant for the higher-ability ones only. They are suggestive that it is the cultural level of tutors which matters the most, not their sociological proximity with their tutees.

1.5.2 Peer group influence

Generally speaking, the results shown in [TABLE 1.5](#) provide clear evidence that tutors matter. They are suggestive that a better selection (or training) of tutors can be a way to improve the efficiency of the intervention. We also explored whether it was possible to improve the program through a more careful constitution of the groups of tutees. As discussed above, the program involves participating in bi-monthly meetings with a specific tutor, but also with a specific group

3. Another possible explanation is that some unobserved characteristics of tutors explain jointly their persistence in the program and the quality of their tutoring. Tutors that are the most involved in the program may be those that are the most likely to persist, but also those that make the program more time consuming for students and, eventually, those that end up having the most negative impact on lower-ability tutees.

TABLE 1.5 – The effect of tutor and assigned peers characteristics on grade 11 achievement

	All	Lower ability	Higher ability
	(I)	(II)	(III)
Panel A: Tutor characteristics			
Female	0.044	0.007	0.001
s.e.	(0.175)	(0.247)	(0.233)
Freshman	0.367***	0.616***	0.316*
s.e.	(0.124)	(0.210)	(0.180)
High-SES	0.622***	0.292	1.019***
s.e.	(0.184)	(0.250)	(0.266)
Obs.	277	130	147
Panel B: Assigned peers characteristics			
Prop. of girls	0.354*	0.010	0.367
s.e.	(0.187)	(0.183)	(0.247)
Prop. of higher-ability	0.263	0.148	0.094
s.e.	(0.249)	(0.274)	(0.350)
Prop. of high-SES	−0.183	−0.555	−0.076
s.e.	(0.313)	(0.452)	(0.461)
Obs.	200	97	103

Panel A shows the results of regressing the average grade obtained at national exams taken at the end of grade 11 by eligible students on three dummies indicating whether the tutor (to whom they were randomly assigned) is a woman, whether the tutor is a ENS freshman and whether the tutor comes from a high SES family background.

Panel B shows the results of regressing the same dependent variable on variables describing the composition of the group of peers to which the individual was randomly assigned (proportion of girls, proportion of higher ability students, proportion of students with high-SES family background).

The sample includes eligible students from both cohorts in panel A, from cohort 2011 only in panel B.

Standard errors clustered at the tutoring-group level are reported in parentheses.

***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

of five other tutees. As discussed above, we introduced some controlled randomness in the design of these groups of peers for the first experimental cohort. Specifically, early September, at the end of the introductory week, we first asked students to form (freely) groups of three persons. In a second step, we paired randomly these 3-person groups in order to form the final list of 6-person groups of tutees. In this set-up, it is possible to look whether students randomly assigned to different elementary 3-person groups of peers obtain different results at the end of their 11th grade.

To implement this test, [TABLE 1.5](#) (panel B) focuses on the first cohort of recipients and shows the result of regressing their performance at the end of the 11th grade on a variable indicating the proportion of girls in the 3-person group of peers with whom they were randomly matched as well as a variable indicating the proportion of higher-ability students and the proportion of students coming from a high-SES family, controlling for the same set of pre-treatment characteristics as in the previous model. We do not find any effect of the proportion of higher-ability or high-SES peers on the performance of recipients at the end of 11th grade, but a positive effect of the proportion of girls. Further explorations suggest that the positive effect of female peers is particularly significant on higher-ability students.

1.6 Discussion

Generally speaking, our experiment suggests that students can volunteer to participate in a program, and persist in this program, even when this program ends up having strongly negative effects on a large fraction of them. To shed light on this paradox, we started in 2016 a survey on former program participants, *i.e.*, 3 to 4 years after their participation in the program. We asked them about their perceptions of the benefits and drawbacks of the program. We also asked them whether and why they decided to persist in (or quit) the program before the end, and about the quality of their interactions with the tutor or with their six-person group of peers. At this stage, we collected 206 responses from former participants, 111 of whom are lower-ability.⁴

One feature of the program is that it does not simply try to improve students' information on

4. Data collection is still underway and the sample size should increase in the near future.

higher education, but also tries to improve their readiness for higher education, which explains the academic contents of the tutoring session. The potential drawback of such a focus is that it may have a depressing effect on some students. Put differently, one reason for why we observe a negative effect on lower-ability volunteers may be that they get the impression that higher education is too difficult for them and that they will never be able to be as strong as their ENS tutors.

For now, however, the results of the post-treatment survey are not really consistent with this assumption. About 68% of respondents actually disagree with the statement that the tutor was difficult to understand (with no difference between lower-ability and higher-ability students). Also, about 84% of respondents agree that the tutor was close to their students (higher-ability: 88%; lower-ability: 80%) and 84% agree that the tutor was positive and encouraging (higher-ability: 86%; lower-ability: 81%). Overall, a very large majority of respondents, including lower-ability ones, had a positive relationship with their tutor.

A second potential explanation for our results is that the program takes too much time and induces students to allocate less time to their school homeworks. As it happens, the program requires a fairly large amount of time. There are three to five 3-hour sessions per term. These sessions took place on Saturday afternoons in Paris (2pm – 5pm) and many students need 1 hour or more to go to Paris. In addition, most tutors used to give specific homeworks to students with the effect of increasing students' workload between sessions.

The post-treatment survey confirms that tutoring sessions were perceived by many participants as too time consuming. 52% of respondents agreed with the statement that the travel time was long (higher-ability: 60%; lower-ability: 46%). 58% of respondents say that they did not do any school home-works on the Saturdays when the sessions took place; 38% of respondents say they had a lot of work to do for the program between the sessions, and 24% say that they had less time to do their school homework because of the program. These proportions are similar for lower-ability and higher-ability students. In 2010, several students actually complained about the amount of time required by the program and the workload was reduced for the second co-

hort.⁵ Among our survey respondents, 98 respondents quit the program at the end of the first year of treatment and the majority explained that the main reason was that it took too much time.

The program was also time-consuming by providing participants with new friends and with new opportunities to spend time with friends. Three or four year after the program, 90% of respondents report that the atmosphere of their tutoring was pleasant, 69% say that they became friends with other students from their tutoring group and 60% say they had kept in touch with former participants from the program. 52% of students agreed with the statement that they spent a lot of time outside of the sessions with the friends they met during the program and 50% mentioned the friendships as one of the reasons why they persisted in the program in grade 12. Overall, our post-treatment survey does not support the assumption that students were discouraged by tutors, but rather that the program took up too much time, be it because of the length of the sessions themselves, the travel time, the between-session homework or the socialization.

1.7 Bibliography

Avery, C. (2009). The effects of college counseling on high-achieving, low-income students: Results of a pilot study with a randomized controlled trial september, 2009. Technical Report 16359, National Bureau of Economic Research.

Avery, C. (2013). Evaluation of the college possible program: Results from a randomized controlled trial. Technical Report 19562, National Bureau of Economic Research.

Bailey, M. J. and Dynarski, S. M. (2011). Gains and gaps: Changing inequality in u.s. college entry and completion. Working Paper 17633, National Bureau of Economic Research.

Bettinger, E. P., Long, B. T., Oreopoulos, P., and Sanbonmatsu, L. (2012). The role of ap-

5. As mentioned above, this reform was not followed by any clear improvement in the effect of the program, although the small size of the second cohort makes it difficult to draw any strong conclusion.

- plication assistance and information in college decisions: Results from the h&r block fafsa experiment*. *The Quarterly Journal of Economics*, 127(3):1205–1242.
- Castleman, B. L. and Page, L. C. (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior & Organization*, 115:144–160.
- Hoxby, C. M. and Avery, C. (2013). Low-income high-achieving students miss out on attending selective colleges.
- Hoxby, C. M. and Turner, S. (2013a). Expanding college opportunities for high-achieving, low income students. *Stanford Institute for Economic Policy Research Discussion Paper*, (12-014).
- Hoxby, C. M. and Turner, S. (2013b). Informing students about their college options: A proposal for broadening the expanding college opportunities project. *The Hamilton Project, Discussion Paper*, 3.
- Roth, A. E. (2008). Deferred acceptance algorithms: History, theory, practice, and open questions. *International Journal of Game Theory*, 36(3-4):537–569.
- Smith, J., Pender, M., and Howell, J. (2013). The full extent of student-college academic undermatch. *Economics of Education Review*, 32:247–261.

1.A Additional tables

TABLE 1.A.1 – Balancing test on the baseline survey

Variable	Control	T – C	s.e.	Obs.
Interested in politics	0.344	+0.008	(0.049)	475
Interested in economics	0.337	+0.024	(0.048)	471
Interested in national news	0.689	+0.005	(0.047)	476
Interested in international news	0.755	+0.020	(0.046)	475
Interested in litterature	0.490	−0.010	(0.052)	474
Interested in history	0.637	−0.004	(0.053)	474
Interested in science and tech.	0.591	−0.008	(0.045)	472
Interested in sports	0.576	+0.088*	(0.050)	476
Interested in social issues	0.639	−0.018	(0.051)	470
Interested in arts	0.534	−0.027	(0.051)	471
Gets help for homework	0.657	+0.004	(0.049)	478
Positive self-assessment	0.613	+0.062	(0.047)	463
Thinks about a job	0.720	−0.043	(0.047)	487
Ambitions 5+ years higher ed.	0.593	−0.051	(0.052)	460
Knows about <i>Grandes Écoles</i>	0.581	−0.058	(0.053)	473
Knows about CPGE	0.563	−0.031	(0.051)	475
Knows about university	0.548	+0.023	(0.053)	475
Knows about IEP	0.357	−0.045	(0.049)	473
Knows about IUT	0.141	+0.047	(0.040)	476
Knows about BTS	0.394	−0.030	(0.052)	476
Gets help for orientation	0.619	+0.032	(0.051)	469
Believes in equal opportunities	0.416	+0.058	(0.054)	453

This table shows differences between the control and the treatment group in the baseline survey completed by volunteer students at the end of grade 10, just before the random draw.

For each variable, the first column gives the average of the control group, the second column gives the difference between the control and the treatment group controlling for cohort, high school and grade 11 major, the third column gives the standard error of the difference and the last column gives the number of non-missing observations for that variable.

TABLE I.A.2 – Balancing test on the sample of volunteer students

Variable	Control	T – C	s.e.	Obs.
Panel A: All volunteer students				
1+ year behing	0.183	+0.041	(0.035)	558
Female	0.622	−0.036	(0.044)	558
High-SES	0.243	+0.019	(0.038)	558
DNB Math score missing	0.072	+0.007	(0.025)	558
Std DNB Math score	0.730	−0.104	(0.077)	518
DNB French missing	0.068	+0.009	(0.025)	558
Std DNB French score	0.557	−0.071	(0.084)	520
Panel B: Lower-ability students				
1+ year behing	0.365	+0.038	(0.065)	280
Female	0.603	−0.043	(0.064)	280
Upperclass	0.262	+0.008	(0.057)	280
DNB Math score missing	0.111	−0.002	(0.043)	280
Std DNB Math score	0.416	−0.096	(0.123)	248
DNB French score missing	0.103	+0.004	(0.043)	280
Std DNB French score	0.209	−0.009	(0.129)	250
Panel C: Higher-ability students				
1+ year behing	0.000	+0.000	(0.000)	278
Female	0.640	−0.013	(0.065)	278
Upperclass	0.224	+0.006	(0.055)	278
DNB Math score missing	0.032	−0.001	(0.023)	278
Std DNB Math score	1.021	−0.035	(0.106)	270
DNB French score missing	0.032	−0.001	(0.023)	278
Std DNB French score	0.881	−0.077	(0.107)	270

This table shows differences between the control and the treatment group on a set of variables from the administrative datasets.

For each variable, the first column gives the average of the control group, the second column gives the difference between the control and the treatment group controlling for cohort, high school and grade 11 major, the third column gives the standard error of the difference and the last column gives the number of non-missing observations for that variable.

TABLE 1.A.3 – The effect of the treatment on performance on high school exit examinations, by cohort

Dependent variable	Cohort 2010				Cohort 2011			
	Grade 11 average score	Average score	Graduation	Honours	Grade 11 average score	Average score	Graduation	Honours
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Panel A: All volunteer students								
Treatment	−0.031	0.024	0.030	0.023	0.004	−0.005	−0.011	−0.034
s.e.	(0.081)	(0.093)	(0.051)	(0.049)	(0.179)	(0.154)	(0.072)	(0.071)
Obs.	378	381	394	394	145	145	153	153
Panel B: Lower-ability								
Treatment	−0.203*	−0.213	−0.064	−0.086	−0.362	−0.238	−0.133	−0.070
s.e.	(0.112)	(0.153)	(0.076)	(0.068)	(0.310)	(0.230)	(0.112)	(0.095)
Obs.	188	191	202	202	70	70	74	74
Panel C: Higher-ability								
Treatment	0.152	0.309**	0.126**	0.129*	0.334	0.148	0.115	−0.027
s.e.	(0.117)	(0.129)	(0.061)	(0.077)	(0.257)	(0.180)	(0.090)	(0.107)
Obs.	190	190	192	192	75	75	79	79
Panel D: Differential impact								
Treatment × higher ability	0.356**	0.522**	0.189**	0.214*	0.695	0.386	0.248*	0.043
s.e.	(0.170)	(0.221)	(0.093)	(0.111)	(0.425)	(0.287)	(0.147)	(0.142)
Obs.	378	381	394	394	145	145	153	153

The table shows the results from reduced-form regressions in which variables measuring performance on high school exit examinations (*baccalauréat*) are regressed on a treatment dummy, using students' gender, pre-treatment ability score and socioeconomic family background as control variables. Columns (I) to (IV) show the results for cohort 2010 and columns (V) to (VIII) for cohort 2011.

Columns (I) and (V) show the estimated effect of the treatment when the dependent variable is the average grade obtained at examinations taken at the end of grade 11 and columns (II) and (VI) when the dependent variable is the average grade across all examinations. Columns (III) and (VII) show the estimated effect when the dependent variable is a dummy indicating high school graduation (average grade is 10/20 or more) and columns (IV) and (VIII) when the dependent variable is a dummy indicating high school graduation with honours (average grade is 12/20 or more).

For each cohort, the first panel refers to the full sample, the second panel to subsample with lower pre-treatment ability score, the third panel to the subsample with higher pre-treatment ability score. Finally, the fourth panel shows the estimated difference in the effect of the treatment between lower and higher ability students. To estimate this difference, we use the full sample and regress the dependent variable on the interaction between a dummy indicating treatment and a dummy indicating high ability, controlling for the treatment dummy as well as for the full set of socio-demographic controls and their interactions with the higher ability dummy.

Standard errors clustered at the class level are reported in parentheses.

** and * denote significance at the 5% and 10% levels, respectively.

TABLE I.A.4 – The effect of the treatment on performance on high school exit examinations, by track

Dependent variable	Non-scientific tracks			Scientific track			
	Grade 11 average score	Average score	Graduation	Honours	Grade 11 average score	Average score	Honours
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
Panel A: All volunteer students							
Treatment	−0.075	−0.186	−0.070	0.010	−0.007	0.107	0.063
s.e.	(0.123)	(0.118)	(0.074)	(0.068)	(0.096)	(0.098)	(0.048)
Obs.	195	196	200	200	328	330	347
Panel B: Lower-ability							
Treatment	−0.210	−0.274**	−0.147	−0.068	−0.285	−0.168	−0.102
s.e.	(0.148)	(0.122)	(0.092)	(0.073)	(0.183)	(0.223)	(0.087)
Obs.	140	141	144	144	118	120	132
Panel C: Higher-ability							
Treatment	0.657**	0.232	0.269*	0.284*	0.109	0.276**	0.057
s.e.	(0.335)	(0.238)	(0.157)	(0.142)	(0.113)	(0.112)	(0.068)
Obs.	55	55	56	56	210	210	215
Panel D: Differential impact							
Treatment × higher ability	0.868***	0.505**	0.416**	0.352**	0.394*	0.444*	0.159
s.e.	(0.319)	(0.236)	(0.171)	(0.136)	(0.219)	(0.257)	(0.109)
Obs.	195	196	200	200	328	330	347

The table shows the results from reduced-form regressions in which variables measuring performance on high school exit examinations (*baccalauréat*) are regressed on a treatment dummy, using students' gender, pre-treatment ability score and socioeconomic family background as control variables. Columns (I) to (IV) show the results for the non-scientific tracks and columns (V) to (VIII) for the scientific track.

Columns (I) and (V) show the estimated effect of the treatment when the dependent variable is the average grade obtained at examinations taken at the end of grade 11 and columns (II) and (VI) when the dependent variable is the average grade across all examinations. Columns (III) and (VII) show the estimated effect when the dependent variable is a dummy indicating high school graduation (average grade is 10/20 or more) and columns (IV) and (VIII) when the dependent variable is a dummy indicating high school graduation with honours (average grade is 12/20 or more).

For each track, the first panel refers to the full sample, the second panel to subsample with lower pre-treatment ability score, the third panel to the subsample with higher pre-treatment ability score. Finally, the fourth panel shows the estimated difference in the effect of the treatment between lower and higher ability students. To estimate this difference, we use the full sample and regress the dependent variable on the interaction between a dummy indicating treatment and a dummy indicating high ability, controlling for the treatment dummy as well as for the full set of socio-demographic controls and their interactions with the higher ability dummy.

Standard errors clustered at the class level are reported in parentheses.

** and * denote significance at the 5% and 10% levels, respectively.

TABLE 1.A.5 – The effect of the treatment on scores at the *Baccalauréat* in the scientific track

Dependent variable	Grade 11			Grade 12, scientific subjects				Grade 12, literary subjects				Final result			
	Average	French (written)	French (oral)	TPE	Average	Math	Phy-Chem	Biology	Average	History	Philosophy	Lang. 1	Average	Pass	Honours
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)	(XIII)	(XIV)	(XV)
Panel A: All volunteer students															
Treatment	−0.087	−0.044	−0.036	−0.086	0.170*	0.061	0.225**	0.162*	−0.081	0.060	0.041	−0.174*	0.102	0.063	0.019
s.e.	(0.096)	(0.092)	(0.101)	(0.101)	(0.097)	(0.093)	(0.096)	(0.093)	(0.067)	(0.095)	(0.095)	(0.103)	(0.102)	(0.051)	(0.054)
Obs.	338	338	338	338	337	337	337	326	338	338	338	338	338	351	351
Panel B: Lower-ability															
Treatment	−0.270	−0.322**	−0.043	−0.185	−0.017	−0.040	0.167	−0.173	−0.238	−0.114	0.023	−0.465**	−0.194	−0.021	−0.121
s.e.	(0.175)	(0.157)	(0.186)	(0.182)	(0.202)	(0.179)	(0.214)	(0.211)	(0.166)	(0.203)	(0.219)	(0.218)	(0.220)	(0.091)	(0.073)
Obs.	122	122	122	122	122	122	122	113	122	122	122	122	122	131	131
Panel C: Higher-ability															
Treatment	0.028	0.117	−0.043	−0.017	0.276***	0.103	0.289**	0.376***	0.032	0.189	0.100	−0.032	0.285**	0.106**	0.105
s.e.	(0.113)	(0.109)	(0.120)	(0.132)	(0.102)	(0.095)	(0.113)	(0.104)	(0.075)	(0.117)	(0.134)	(0.112)	(0.114)	(0.053)	(0.066)
Obs.	216	216	216	216	215	215	215	213	216	216	216	216	216	220	220
Panel D: Differential impact															
Treatment × higher ability	0.298	0.438**	0.001	0.168	0.293	0.143	0.122	0.549**	0.270	0.302	0.078	0.433*	0.479*	0.128	0.226**
s.e.	(0.209)	(0.182)	(0.216)	(0.222)	(0.217)	(0.186)	(0.245)	(0.233)	(0.191)	(0.236)	(0.277)	(0.244)	(0.254)	(0.096)	(0.097)
Obs.	338	338	338	338	337	337	337	326	338	338	338	338	338	351	351

The sample includes volunteer students from cohorts 2010 and 2011 who specialized in science in grades 11 and 12.

The table shows the results from reduced-form regressions in which variables measuring performance on high school exit examinations (*baccalauréat*) are regressed on a treatment dummy, using students' gender, pre-treatment ability score and socioeconomic family background as control variables.

Column (I) shows the estimated effect of the treatment when the dependent variable is the average grade obtained at examinations taken at the end of grade 11; columns (II), (III) and (IV) show the effect for each subject taken that year. Column (V) shows the effect when the dependent variable is the average grade obtained at examinations taken at the end of grade 12 in scientific subjects; columns (VI), (VII) and (VIII) show the effect for each scientific subject taken that year. Column (IX) shows the effect when the dependent variable is the average grade obtained at examinations taken at the end of grade 12 in literary subjects; columns (X), (XI) and (XII) show the effect for each literary subject taken that year. Column (XIII) shows the effect when the dependent variable is the average grade across all examinations, column (XIV) when the dependent variable is a dummy indicating high school graduation (average grade is 10/20 or more) and column (XV) when the dependent variable is a dummy indicating high school graduation with honours (average grade is 12/20 or more).

The first panel refers to the full sample, the second panel to subsample with lower pre-treatment ability score, the third panel to the subsample with higher pre-treatment ability score. Finally, the fourth panel shows the estimated difference in the effect of the treatment between lower and higher ability students. To estimate this difference, we use the full sample and regress the dependent variable on the interaction between a dummy indicating treatment and a dummy indicating high ability, controlling for the treatment dummy as well as for the full set of socio-demographic controls and their interactions with the higher ability dummy. Standard errors clustered at the class level are reported in parentheses.

***, **, and * denote significance at the 1%, 5% and 10% levels, respectively.

TABLE I.A.6 – The effect of the treatment on scores at the *Baccalauréat* in the non scientific tracks

Dependent variable	Grade 11			Grade 12, scientific subjects					Grade 12, literary subjects					Final result		
	Average	French (written)	French (oral)	TPE	Math	Social Sc.	Sciences	Average	History	Philosophy	Lang. 1	Average	Pass	Honours		
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)	(XIII)	(XIV)		
Panel A: All volunteer students																
Treatment	-0.098	-0.110	-0.061	-0.137	-0.064	-0.115	0.058	-0.088	-0.216	0.065	-0.194	-0.164	-0.072	0.009		
s.e.	(0.117)	(0.144)	(0.130)	(0.148)	(0.119)	(0.241)	(0.137)	(0.079)	(0.132)	(0.160)	(0.127)	(0.117)	(0.074)	(0.068)		
Obs.	199	199	199	199	184	110	198	196	199	199	196	199	200	200		
Panel B: Lower-ability																
Treatment	-0.230*	-0.202	-0.125	-0.359*	-0.131	-0.014	0.086	-0.107	-0.337**	0.082	-0.253*	-0.277**	-0.159	-0.082		
s.e.	(0.124)	(0.158)	(0.153)	(0.182)	(0.159)	(0.317)	(0.161)	(0.082)	(0.152)	(0.205)	(0.150)	(0.121)	(0.096)	(0.066)		
Obs.	142	142	142	142	129	71	141	139	142	142	139	142	142	142		
Panel C: Higher-ability																
Treatment	0.325	0.016	0.209	0.430	0.127	-0.243	0.482*	-0.229	-0.096	-0.137	-0.322	0.067	0.193	0.196		
s.e.	(0.294)	(0.325)	(0.279)	(0.329)	(0.221)	(0.441)	(0.280)	(0.208)	(0.282)	(0.299)	(0.364)	(0.206)	(0.117)	(0.174)		
Obs.	57	57	57	57	55	39	57	57	57	57	57	57	58	58		
Panel D: Differential impact																
Treatment × higher ability	0.555**	0.218	0.334	0.789**	0.258	-0.229	0.397	-0.123	0.241	-0.219	-0.069	0.344	0.352**	0.278*		
s.e.	(0.266)	(0.303)	(0.280)	(0.326)	(0.263)	(0.489)	(0.284)	(0.190)	(0.275)	(0.310)	(0.352)	(0.208)	(0.145)	(0.154)		
Obs.	199	199	199	199	184	110	198	196	199	199	196	199	200	200		

The sample includes volunteer students from cohorts 2010 and 2011 who specialized in humanities in grades 11 and 12.

The table shows the results from reduced-form regressions in which variables measuring performance on high school exit examinations (*baccalauréat*) are regressed on a treatment dummy, using students' gender, pre-treatment ability score and socioeconomic family background as control variables.

Column (I) shows the estimated effect of the treatment when the dependent variable is the average grade obtained at examinations taken at the end of grade 11; columns (II), (III) and (IV) show the effect for each subject taken that year. Column (V) shows the effect when the dependent variable is the average grade obtained at examinations taken at the end of grade 12 in scientific subjects; columns (VI) and (VII) show the effect for each scientific subject taken that year. Column (VIII) shows the effect when the dependent variable is the average grade obtained at examinations taken at the end of grade 12 in literary subjects; columns (X), (XI) and (XII) show the effect for each literary subject taken that year. Column (XIII) shows the effect when the dependent variable is the average grade across all examinations, column (XIV) when the dependent variable is a dummy indicating high school graduation (average grade is 10/20 or more) and column (XV) when the dependent variable is a dummy indicating high school graduation with honours (average grade is 12/20 or more).

The first panel refers to the full sample, the second panel to subsample with lower pre-treatment ability score, the third panel to the subsample with higher pre-treatment ability score. Finally, the fourth panel shows the estimated difference in the effect of the treatment between lower and higher ability students. To estimate this difference, we use the full sample and regress the dependent variable on the interaction between a dummy indicating treatment and a dummy indicating high ability, controlling for the treatment dummy as well as for the full set of socio-demographic controls and their interactions with the higher ability dummy.

Standard errors clustered at the class level are reported in parentheses.

***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Chapitre 2

Persistent Classmates : How Familiarity with Peers Protects from Disruptive School Transitions

Ce chapitre a été écrit avec Son Thierry Ly.

Il a été publié en tant que document de travail de l'École d'économie de Paris (numéro 2013-21).

Nous remercions Éric Maurin, Julie Berry Cullen, Luc Behaghel, Gordon Dahl, Victor Lavy, Thomas Piketty, Corinne Prost, Gwenaél Roudaut, Camille Terrier et Margaux Vinez pour leurs commentaires et suggestions. Nous remercions également les participants des conférences SOLE (Arlington, 2014) et IWAEE (Catanzaro, 2013), du séminaire d'économie appliquée de l'École d'économie de Paris ainsi que du séminaire interne du Crest. Nous sommes également reconnaissants à la Direction de l'évaluation, de la prospective et de la performance au Ministère de l'Éducation nationale et en particulier à Cédric Afsa qui a facilité notre accès aux bases de données.

Cette recherche a bénéficié d'un soutien financier du CEPREMAP.

2.1 Introduction

If peer effects exist, policies that influence the allocation of individuals with their peers could improve welfare. This idea has aroused a huge amount of interest by economists, despite the great empirical challenges raised by endogenous sorting. For example, many papers investigate the role of neighborhood (Goux and Maurin, 2007; Kling et al., 2007) and school composition (Hoxby, 2000; Angrist and Lang, 2004; Cullen et al., 2006; Lavy et al., 2012b) on students' outcomes. The literature is much less extensive when it comes to estimating peer effects within classes, although most students' interactions are likely to occur at this level. It is also unfortunate from a policy point of view, because school administrators have much more leeway in setting up classes than policymakers have in influencing neighborhood and school choice. The main studies on the subject are based on experimental data either in primary schools in developing countries (Duflo et al., 2011) or in colleges in developed countries (Carrell et al., 2011). Evidence based on observational data is rare (see e.g. Lavy et al., 2012a; Fruehwirth, 2013), since it requires both rich data at class level and conclusive natural experiments that are rarely available.

In this paper, we identify classroom peer effects by using natural experiments aroused by the specific institutional features of student allocation across classes, within schools, in the first year of high school in France (10th grade).¹ By definition, a high school principal does not know her first-year students before the beginning of the school year. As she has to allocate them across classes before that time, she has only a finite set of information observed in their registration files to go on, including e.g. gender, socioeconomic status (SES), middle school and 9th grade class of origin, the list of optional courses and scores obtained in their grade 9 class. Using a unique administrative dataset, we are able to observe almost all characteristics observed by principals about their cohort of freshmen.

Usually, the registration files a given principal has to consider in a given school year are all different from one another. But there are rare cases where she gets two freshmen registration files that are exactly or very similar with regard to *all these characteristics* (denoted "similar-file" or SF students). For example, these two students, call them Aurélien and Benoit, are both low-

1. In France, secondary education consists of two blocks: middle school (grades 6 to 9) and high school (grades 10 to 12). High schools are generally separate from middle schools.

SES boys coming from the same 9th grade class from the same middle school, who got very similar scores in grade 9, and ask for the same optional courses in grade 10. If the high school principal decides to separate them across two 10th grade classes X and Y, the key intuition is that *the choice of assigning Aurélien to class X and Benoit to class Y or the other way around should be as good as random*, because she does not have any additional relevant information to distinguish between them (see [FIGURE 2.2](#)). Note that we do not assume that the decision to separate the students or to keep them in the same class is exogenous, but only that the assignment is random in case she decides to separate them. We are able to provide strong evidence supporting this assumption, using the anonymous scores obtained at the national exam taken by students at the end of 9th grade, *which is unobserved by principals*.

Our estimation strategy is as follows. First, we restrict our analysis to students who have a similar-file mate who ends up in a different class in the same high school. This is the case for only 0.8 percent of the population of high school freshmen, leaving us with a sample of 28,053 students over the 2004-2011 period. Then, our strategy is to compare a student with his or her similar-file mate only. We do this by controlling in all regressions for a single fixed effect accounting altogether for freshmen's high school, cohort, middle school and class of origin, and all other characteristics observed in registration files. Finally, we investigate how the outcome gap between two similar-file students relates to differences in the characteristics of their 10th grade classes. In other words, the random assignment of Aurélien and Benoit between classes X and Y can be seen as a given lottery or quasi-experiment, and the whole set of comparable lotteries happening in French high school between 2004 and 2011 allows us to estimate the effect of a number of classroom environment dimensions.

Note that our strategy differs fundamentally from matching strategies. In a matching strategy, econometricians use observed variables to match students in order to reduce the confounding bias, though the bias does not vanish to zero because of unobserved variables that are likely to be correlated to potential outcomes. In our case, the set of variables that we observe is almost exactly identical to the set of variables that determine the class assignment by the high school principal. This would not be the case if, for instance, high school principals had a personal knowledge of the students, or if the students had to write application essays. We provide strong

evidence that class assignment is indeed random conditional to our fixed effect.

Common measures of peer characteristics are considered in our main regressions, such as peer ability, gender and socioeconomic status (SES). Yet surprisingly, the most robust effect that emerges comes from the number of *persistent classmates* (PC) a student gets, *i.e.* classmates who were already in the freshman's class in the last year of middle school. Not only does the number of PCs significantly reduce the risk of repeating freshman year, but by contrast with other measures of peer characteristics, the effect also endures in the long run and is associated with differences in graduation rates at the end of high school.

The second part of this paper sets out to understand why the presence of these persistent classmates generates positive spillovers. Although the number of PCs might capture some omitted class characteristics associated with peer ability for instance, our investigations suggest that students benefit from having more persistent classmates only because of a *familiarity* mechanism, *i.e.* because they know each other well. Three findings lead us to this conclusion. First, the estimates are extremely robust to the inclusion or not of controls for the other classmates' characteristics (ability, gender and SES). Second, we find that the PC effect is highly heterogeneous and mainly driven by low-achieving, low-SES students. Also, the effect seems slightly stronger when these students are suddenly more exposed to high-SES students. This is consistent with our interpretation, as being surrounded by familiar faces should matter more when the transition to high school is highly disruptive. Third, these students at risk of underachievement in high school are not more impacted by their high- than their low-achieving persistent classmates, which would be expected if the PC effect was driven by their higher unobserved ability. Robustness checks are provided for our main results.

This study makes three important contributions to the literature. First, it sheds light on the ongoing debate on the complexity of peer effects. While some recent studies offer an insight into the role of social networks during school transitions (see *e.g.* Lavy and Sand, 2012), we use natural experiments that provide a stronger identification of the impact of classmates' characteristics, including their social links. In keeping with Foster (2006), our results also take issue with popular belief that agents are more influenced by their friends than by other peers (see also

Halliday and Kwak, 2012). In particular, students could be influenced by former classmates that are not friends, though there was no theoretical reason to expect classmate persistence to have positive effects.² As a matter of fact, recent results found by De Giorgi and Pellizzari (2013) on Bocconi University were suggesting the opposite, as they find a decrease in performance for undergraduate students that are assigned more often together across classes. Our results show that former peers generate positive spillovers when reassigned together in the context of school transitions, emphasizing how peers may not have the same effect depending on the timing and contexts.

Second, our findings need to be considered in relation to the strand of literature on the impact of mobility across environments. Several papers find that policies that enhance neighborhood or school choice, or expand students' access to high-performing schools have been unexpectedly inefficient in improving students' educational outcomes (Angrist and Lang, 2004; Cullen et al., 2006; Kling et al., 2007). In line with other recent works (Lavy and Sand, 2012; Gibbons et al., 2013), this paper suggests that these results could be due to the disruption caused to a student's environment by such policies.

The third contribution is more policy-oriented. Policy recommendations to improve achievement in high school are particularly relevant, in view of the issues at stake. In many countries, formal tracking is implemented in high school, such that short-term low achievement in the first year may end up with mismatched enrollment in low-skill tracks. In addition, the start of high school is often simultaneous with the end of compulsory schooling, meaning that under-achievement may lead to drop-out at that stage compared to previous stages. As our analysis show, principals could substantially raise the achievement in high school of low-ability students

2. On the one hand, former classmates may be friends, and recent evidence suggest that friends may have a positive effect on well-being and achievement (Calvò-Armengol et al., 2009; Lavy and Sand, 2012). Yet former classmates may also simply be peers with whom it is easier to talk during the early weeks, to sit next to in the classroom or to ask for help, thus making it easier to adapt to higher academic expectations and less supervision from teachers. Even without friendship bonds, familiarity within the classroom could therefore reduce anxiety, prevent social isolation and foster a student's sense of belonging in the new school and class. On the other hand, former classmates could prevent students from socializing with new peers, or be conducive to bad behavior in the classroom if disruptive students stay together. Former classmates may also be enemies rather than friends, and their presence could be detrimental to welfare and achievement. Mora and Oreopoulos (2011); Lavy and Sand (2012) show that "non-reciprocal friends" (peers that consider you as a friend while you do not, or vice versa) seem to have no or negative effects on outcomes.

by assigning them in freshman year with some familiar classmates. Our estimates suggest that their risk of repeating freshman year could be reduced by 4.5 percentage points, and their graduation rate raised by the same amount. This simple recommendation on class composition may thus improve their performance by around 13 percent at no cost, while highly expensive policies usually target this population of students at risk. Moreover, moving these students across classes based on their former networks is not a zero-sum game, in contrast to their ability or gender. Although a high-ability student or a female student might be of benefit to everyone,³ grouping together freshmen from the same class should not affect freshmen from other classes.

Section 2.2 describes the institutional context and the data. Section 2.3 describes the identification strategy. We present the results and discuss the distribution of the effect and its mechanisms in section 2.4. Robustness checks are then provided in section 2.5. Section 2.6 discusses the implications of our results and concludes.

2.2 Institutional context and data

2.2.1 The high school curriculum in France

2.2.1.1 Enrolling in general high schools

By the end of middle school (grades 6 to 9), students apply for either vocational or general studies, with the approval of middle school teachers. Around two-thirds of 9th grade students opt for the general track, in which case they apply to general high schools in their district.⁴ Rules of admission then differ by school district and year, but they usually depend on the students' home address, socioeconomic status and school performance (9th grade scores). Allocation is over by the end of June and high school administrations receive the registration files on their future 10th graders in the first week of July.

3. Carrell et al. (2011) built an algorithm designed to optimize peer effects and failed to do so partly for this reason.

4. Students opting for the vocational track have to choose a specialty, and vocational high schools usually have places in only one or two classes per specialty. We have therefore decided to exclude vocational high schools from this study, since class composition is highly constrained and is not really policy-relevant in these schools.

At the same time, 9th grade students take national anonymous exams in the end of June in three core subjects: mathematics, French and history-geography. These exams are not graded by teachers from the student's middle school, but externally (with scores between 0 and 40). The resulting *anonymous scores* are combined with *continuous assessment scores*, *i.e.* scores obtained in 9th grade in all courses and graded by the students' own teachers (between 0 and 20). The anonymous scores and continuous assessment scores are combined to compute a total score that determines whether they pass the middle school graduation diploma (*Diplôme national du brevet* or DNB hereinafter).⁵

The anonymous scores are only available in mid-July. By that time, students have completed their administrative registration for high school and class compositions are already determined. In addition, these scores are not sent to the high school during the summer⁶. Therefore, the principals assign freshmen to grade 10 classes without knowing their anonymous scores, and only having their continuous assessment scores.

2.2.1.2 The curriculum in general high schools

In France, freshman year marks a difficult milestone for students attending general high schools. The average ability of peers raises suddenly as one third of students, usually the lowest achieving ones, has enrolled in vocational studies after middle school. This may not only increase teachers' expectations, but it might also affect negatively students' self-esteem. Students' percentile rank within their class drops from 64 to 52 in average between grade 9 and 10. Naturally, this change hits students asymmetrically. While students in the top half of the ability distribution only fall from the 78th to the 71st percentile rank, the other half drops from the 49th to the 32nd percentile rank.⁷ Students also undergo a shock in the sociocultural dimension, as reflected by the share of high-SES that increases from 22 percent in middle school to 30 percent in general high school.

5. Note that students do not need to pass to go on to high school.

6. Some students do inform the high school of their results in the anonymous exams once they receive them (although this is not a requirement), but informal discussions we have had with some high school administrations suggest that this hardly ever happens. In any case, principals do not have these scores for all students, so they are highly unlikely to use them to assign students across classes.

7. These figures are computed using the anonymous exam scores at the DNB exam.

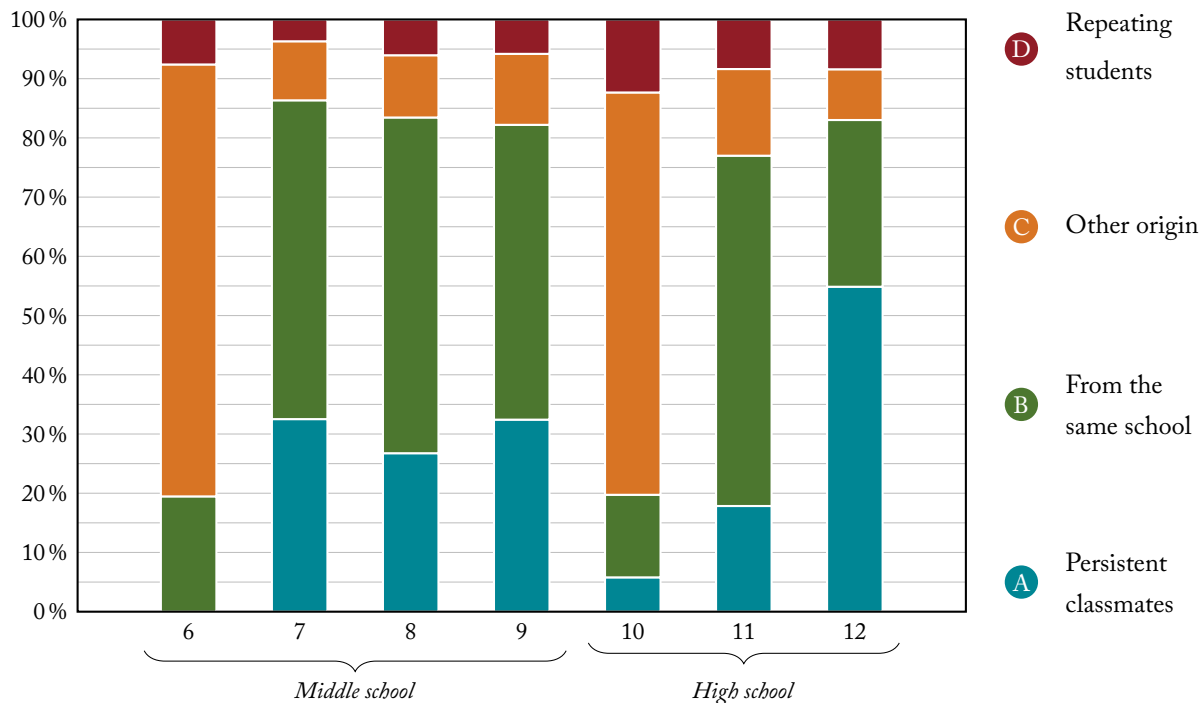
At the same time, the school change triggers an intense disruption in students' networks, as illustrated by the amount of *persistent classmates* in their class. **FIGURE 2.1** plots the typical student's class composition for each grade. As a benchmark, the share of persistent classmates remains fairly constant throughout middle school at around 30 percent.⁸ Yet in grade 10, the number of PCs drops dramatically. Only 5 percent of their classmates come from the same class and 20 percent from the same middle school. Assuming that students rarely know the students from other middle schools, this means that students do not know at least 80 percent of their classmates at the beginning of the year. This figure decreases again in the subsequent grades, coming to 45 percent in grade 12 due to the partial carryover of major-specific classes from grade 11.

It turns out that these disruptions happen precisely at a time when achievement is highly determinant for long-term outcomes. By the end of the year, students have to apply for a major that will determine their 11th and 12th grade courses, their *baccalauréat* (high school graduation exam) specialty, and the university tracks they will be able to apply for at the end of high school. First, students have to opt for the academic or technological track, the former being historically more prestigious with harder, more academic courses. If students are not accepted for any of the majors they apply for, they can opt for an alternative major suggested by teachers, if any. Otherwise, they have to repeat grade 10 with a view to applying again the following year.⁹ As a result, repetition rate is exceptionally high in the first year of high school (10 percent compared to 5 percent in average in middle school).

Lastly, high school ends at grade 12 with the *baccalauréat* exam. This high school graduation exam includes anonymous tests in different subjects depending on the student's major, and is almost entirely graded by teachers outside the student's high school. Passing the *baccalauréat* is required in most higher education tracks, and it is sufficient to access most university tracks.

8. In grade 6, we are only able to identify students from the same elementary school, as we do not have any information on the classes in grade 5.

9. Students not allowed to move up to the next grade may appeal the decision to a committee external to the school, whose decision is final. Those refused permission to enroll in the major of their choice may appeal the decision to the principal or even negotiate with a different high school. In any case, the final decision to award a student a place on a given major rests with the principal of the high school attended in 11th grade. The principals we met reported that very few students in each year actually go against their teachers' advice.



This graph shows the typical composition of a student's classroom when he or she enters a given grade $g \in \{6, \dots, 12\}$. For each grade g , the sample consists of students entering that grade for the first time (*i.e.* repeating students are excluded) for cohorts 2000 to 2008 (when they never repeat, students from cohort 2000 enter grade 6 in 2000, grade 7 in 2001, etc.).

READING: Among the classmates of a student entering grade 9, 32 percent are persistent classmates from 8th grade, $82 - 32 = 50$ percent are classmates who were in different classes from the same school in 8th grade, $94 - 82 = 12$ percent are non-repeating students who were in a different school in 8th grade, and 6 percent are students repeating 9th grade.

NOTES ABOUT THE DATA: Cohorts 2000 and 2001 are missing for grades 6 and 7; cohort 2008 missing for grade 12. We only know the school attended in grade 5, not the classroom, therefore we cannot distinguish between persistent classmates and former schoolmates (stacks A and B) in grade 6.

Note that during middle school, the average share of persistent classmates is around 30 percent. This is due to the fact that classes are generally reshuffled from one year to another.

FIGURE 2.1 – Composition of the typical classroom from a non-repeating student's point of view

2.2.2 The class-assignment mechanism

In France, students are assigned to the same class for all subjects for the entire school year. Classmates therefore have even more potential influence over each other's outcomes, as they spend most of the day together throughout the school year. In practice, classes are assigned in early July immediately following student registration for high school, and two months before the start of the school year in September. Classes are assigned entirely by hand, without the aid of computer algorithms. This process is non-random, even for first year students. However, unlike with the other grades, high school principals do not know the students personally when they assign them to classes. Consequently, high school principals rely solely on the set of formal registration data given in the students' files and observable, for the most part, in our dataset.¹⁰

First, principals look at the options chosen by students. While most courses are part of the common core curriculum and are the same for all (*e.g.* mathematics and French), students have certain subjects to choose from such as which foreign language they prefer to study (*e.g.* English or Spanish) and some additional optional courses (*e.g.* Latin and ancient Greek). Students who take the same options are often grouped in the same class, for the sake of convenience when timetabling classes.

Conditional on students' options, school principals generally (but not necessarily) try to balance classes in terms of gender and ability.¹¹ They rely on the formal data contained in students' personal registration files: personal details on the students and their families (mainly gender, age and parents' occupations), scores obtained in 9th grade subjects (between 0 and 20) and 9th grade teachers' comments. These short comments are not written for principals but for

10. High schools are usually separate from middle schools. However, 16 percent of French students attend schools that cover the entire secondary curriculum (mostly private schools). In these schools, principals might know 10th grade students coming in from their own middle school. Nonetheless, middle school and high school still have separate deputy heads to whom principals generally delegate class composition. These deputy heads do not necessarily coordinate over class assignment of 10th grade students, such that the high school deputy heads may well not use any more information than the registration file. This is supported by our exogeneity test (see section 2.3.1), which suggests that students are conditionally randomly assigned even in this case. We have therefore chosen to keep students from these schools in our sample. Taking them out of the sample has virtually no impact on the results.

11. There is no legal requirement to do so, but the 1975 Haby Act that made middle schools comprehensive established a tacit rule for schools to favor within-class heterogeneity. Besides, principals probably want to avoid putting all low-achieving students together in one class that consequently risks being unruly.

parents, to assess student performance and behavior in each subject. For example, teachers may write that they are satisfied with the student's effort and participation in class, or that the student talks too much with his or her classmates in class (without naming names). These reports do not include recommendations to high school principals such as, "Do not put these two students together". Only on the rarest of occasions would such advice be given to principals, and then via an informal channel.

Unlike with other grades where principals know their students, they cannot count on any personal knowledge of them such as motivation or emotional resilience.¹²

In addition, in French high schools, neither parents nor students can ask for placement in specific classes or with friends. As a rule, families do not liaise directly with principals over class assignment.¹³ Instead, tactics to get children assigned to a better class mainly take the form of choosing specific options. In particular, families may encourage their children to take "elite" options (*e.g.* German as a first foreign language, or Latin) to get them assigned to a better class. This has no impact on our identification since we only compare students taking the same options. Lastly, students are only notified of their class assignment the week before the first day of school, and they are not allowed to change it.

There are good reasons to believe that principals do not use all the detailed formal data they have on students to assign them to a class. As revealed by the class assignment sessions we attended, simply allocating classes on the basis of options is already complicated and time-consuming enough as it is. Again, they have to do it by hand and take a large number of constraints into account, while a host of other tasks are pending both to wind up the current school year and prepare for the new one. Therefore, if two freshmen's registration files look broadly similar, principals are highly unlikely to spend time studying their characteristics to try to find some minor detail to differentiate between them. In particular, principals do not telephone families or middle school principals to get further information on students. In practice, then, two

12. With other grades, they might, for example, separate two friends who are disrupting lessons, or place a fragile student with his or her friends for emotional support.

13. They do so only in very special cases, such as where car sharing needs to be organized for students in rural areas.

10th grade students do not need to be exactly identical on paper to be deemed indistinguishable during the class assignment process. Section 2.3 provides empirical evidence in support of this field observation.

2.2.3 Data

2.2.3.1 Datasets

The empirical analysis is based on two administrative datasets from the French Ministry of Education.

- *Administrative registration records*: for all students enrolled in French public and publicly-funded private middle and high schools from 2001 to 2012. This dataset contains students' personal details (*e.g.* date and region of birth, gender and parents' occupation) and information on their education: in particular grade, school and class attended, options taken, grade and school attended in $t - 1$ (but not the class attended in $t - 1$).
- *Examination records*: for all students from 2004 to 2011. This dataset contains personal details and informal scores in the 9th grade DNB (both the anonymous exam and continuous assessment scores) and 12th grade *baccalauréat* exams.

These datasets are exhaustive and the variables we make use of are well reported for almost 100 percent of the population. Unfortunately, students do not have personal identification numbers so that they can be tracked through the different datasets. Yet for each 10th grade student, we need to know at least which class they attended in 9th grade, their grade in $t + 1$ (repeating 10th grade or moving to 11th grade) and chosen major if they do move to 11th grade. We also have to match the administrative and the examination records.

In order to find this missing information, we use a matching procedure taking the students' personal details in each dataset. The procedure is based mainly on date and region of birth, gender, grade and school attended in years t and $t - 1$. We manage to match 9th grade class for 94 percent and DNB exam scores for 81 percent of new 10th grade students. The remaining students either had no match in the auxiliary dataset (60 percent of occurrences) or multiple

matches (40 percent of occurrences).¹⁴ The online appendix provides further details on the matching procedure. In the rest of the paper, all regressions include controls for the share of missing observations in the class, although they do not change the estimates.

Our identification compares the set of information on students observed in our dataset to the information observed by principals in their registration files at the time of class allocation. So it is useful at this stage to summarize which variable is observed by whom:

- *Covariates observed by both the principal and the econometrician:* Date of birth, city of residence, gender, parents' occupation, foreign languages and options chosen, 9th grade continuous assessment scores in all subjects, middle school and 9th grade class. We also observe a numerical measure of student behavior as graded by the student's head teacher; this information is missing for the first two cohorts (out of eight) and it will therefore be used only for robustness checks.
- *Covariates observed by the principal, but not the econometrician:* Students' first and last name (from which, in particular, ethnicity could be inferred), and exact home address. The principal also observes the 9th grade teachers' written comments, which may inform of behavioral issues. Again, these comments are very short (one sentence from each teacher), written for parents, and do not include information about the relationships with specific students.
- *Covariates observed by the econometrician, but not by the principal:* Anonymous DNB exam scores.

Most information observed by the principal is thus contained in the dataset. Although we do not observe the teachers' written comments, we do observe a behavioral score for three-quarters of the sample, which contains precisely the information we expect the principals to infer from the written comment¹⁵. As we will show, the anonymous DNB exam scores are key in this study, since they allow us to test our main identification assumption.

14. A multiple match means that two students are found in the same school $\times$ grade $\times$ year with the same date of birth, gender, etc.). This may occur only randomly and is not likely to bias our results.

15. Note that it would be hard, in any case, to work directly with the written comments even if we could observe them. If we were to do so, we would try to build a score to summarize the information contained in the comment, which is the purpose of this behavioral score.

Descriptive statistics are presented in [TABLE 2.1](#) for the entire population of 10th grade students (column I). In particular, it is interesting to note that the average freshman has 1.7 persistent classmates out of the 8.3 former classmates enrolled in their high school.¹⁶

2.3 Identification

The identification strategy used in this paper is based on a quasi-experimental setting. Since principals do not know first-year students personally, principals cannot easily distinguish between two students who are similar "on paper", *i.e.* who exhibit identical or very close registration files. If such students are separated into different classes, it is credible to assume that they were randomly assigned to their classes. Therefore, classroom peer effects are identified by sticking to comparisons between students coming to a given high school with the same observable characteristics, but ending up in different classes. Such students form what we call "*similar-file*" groups (or SF groups), and all comparisons throughout the paper are made between students belonging to the same SF group.

[FIGURE 2.2](#) illustrates this approach, where students A and B have close observable characteristics. If the high school principal has to split them between classes X and Y, we assume that the decision between assigning A to X and B to Y (case 1) or the reverse (case 2) is as good as random. Therefore, the differences between students A and B's classes (*e.g.* characteristics of classmates C-D-E compared to F-G-H) are uncorrelated with differences in individual unobserved factors of achievement, allowing for causal inference of peer effects. Classroom peer effects can be estimated by examining the correlations between students A and B's gap in outcome and differences in their class characteristics.

Formally, we estimate the following model using OLS:¹⁷

$$y_{igc} = \alpha_g + \beta \cdot C_{igc} + \epsilon_{igc} \quad (2.1)$$

16. The ratio is roughly equal to 5, the average number of classes in high schools, suggesting that principals do not try to group students having the same class of origin when allocating them among 10th grade classes.

17. Although our estimation strategy is similar in spirit to exact-matching methods, we choose not to use matching estimation as the regressors examined in this paper are not binary. To our knowledge, the literature is very poor when it comes to the estimation of average causal effects of multi-valued treatments by propensity score or exact matching methods (see [Imbens, 2000](#), from this point of view).

TABLE 2.I – Descriptive statistics on students' characteristics

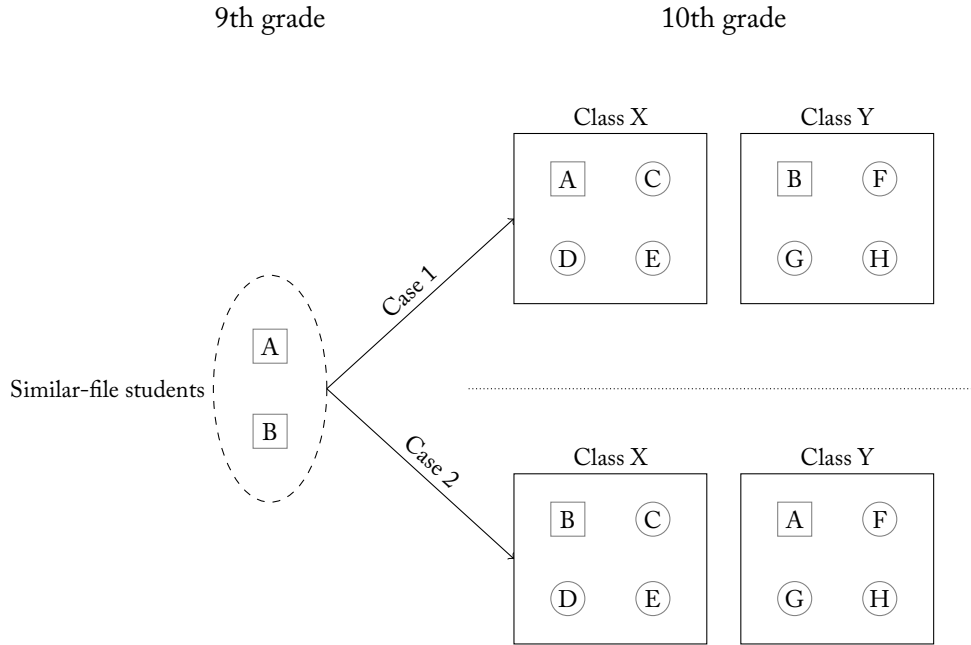
	Population ¹	SF sample	At risk
	(I)	(II)	(III)
Girl	0.547 (0.498)	0.618 (0.486)	0.609 (0.488)
High-SES	0.301 (0.459)	0.301 (0.459)	0.000 (0.000)
High quality optional course	0.145 (0.352)	0.093 (0.291)	0.021 (0.143)
DNB national exam score	23.935 ² (5.077)	25.140 (5.551)	20.114 (3.879)
Normalized DNB national exam score	0.000 ² (1.000)	0.246 (1.099)	−0.757 (0.762)
Had repeated at least once before grade	0.147 (0.354)	0.037 (0.188)	0.101 (0.301)
Repeats 10th grade	0.151 (0.358)	0.150 (0.357)	0.359 (0.480)
Attrition (drop out or unmatched in panel)	0.083 (0.276)	0.040 (0.197)	0.078 (0.269)
Academic major in grade 11	0.590 (0.492)	0.694 (0.461)	0.303 (0.460)
Technological major in grade 11	0.175 (0.380)	0.117 (0.321)	0.259 (0.438)
Takes <i>Bac</i> on schedule	0.658 (0.474)	0.737 (0.441)	0.482 (0.500)
Graduates high school	0.529 (0.499)	0.630 (0.483)	0.314 (0.464)
Number of PC in 10th grade class	1.718 ³ (2.528)	1.999 (2.256)	1.503 (1.760)
Number of former classmates in high school	8.327 ³ (6.466)	12.680 (5.924)	10.934 (4.936)
<i>N</i>	3,589,710	28,053	8,981

Standard deviations are reported in parentheses.

¹ The population is made of the 10th grade students in the general track coming from the 9th grade (*i.e.* excluding repeating students). The general track typically contains more girls and more high-SES students.

² Only students whose DNB score is known: $N = 2,897,986$.

³ Only students whose 9th grade class is known: $N = 3,381,271$.



Consider two "similar-file" students A and B, *i.e.* two first-year students who share the same characteristics for all the information included in the registration files. Our key assumption is that they are not distinguished by the high school principal during class assignment, because she does not know personally her first-year students at that time.

The principal has two options when she allocates these students: she can either assign them to the same class or separate them into two classes X and Y. We do not compare these two scenarios. We focus exclusively on the scenario where the students are separated. In this scenario, there are two possible cases: student A is assigned to class X and student B is assigned to class Y (case 1) or student A is assigned to class Y and student B is assigned to class X.

We argue that the choice between case 1 and case 2 is random and can be seen as a lottery that affects the two students' social environments in 10th grade. For instance, student A will be with students C, D and E in case 1, students F, G and H in case 2.

Differences in outcomes between students A and B can be credibly attributed to differences in classroom characteristics.

FIGURE 2.2 – Class assignment of similar-file students

where y_{igc} denotes high school outcomes for student i , assigned to 10th grade class c and belonging to SF group g , as in model (2.3), C_{igc} is a vector of class characteristics (*e.g.* peer ability, female share, or i 's number of persistent classmate), and ϵ_{igc} captures individual unobserved factors of achievement. α_g is the SF group fixed effect that restricts the analysis to comparisons within groups of students with similar registration files and is the key to identification. Indeed, β captures the peer effects of interest under the key assumption that 10th grade class characteristics C_{igc} are not correlated with ϵ_{igc} *conditional on* g :

$$C_{igc} \perp \epsilon_{igc} | g \quad i.e. \quad \text{Cov}(C_{igc}, \epsilon_{igc} | g) = 0 \quad (2.2)$$

By controlling for SF group fixed effects, model (2.1) estimates β parameters only by comparing *separated similar-file students* with each other, as long as their classrooms differ on C_{igc} dimensions (in particular, we do not compare students who are separated versus students who end up in the same class).

In what follows, we describe how we define SF groups before providing empirical evidence supporting assumption (2.2).

2.3.1 Definition of similar-file groups

The natural experiments consists in students who enrolled in a given high school and year with the same or very similar observable characteristics from the principal's perspective, but assigned to different classrooms. These groups of students, denoted $g \in \{1, \dots, G\}$ and called SF groups, are defined using all variables that we observe on first-year students' registration files. We consider that students belong to the same "similar-file" groups only if they come from the same 9th grade class in middle school; enroll in the same high school in the same year; select the same options (*i.e.* same foreign language and optional courses); share the same gender, age¹⁸ and social background (low- or high-SES) based on father's occupation; belong to the same quintile of average 9th grade continuous assessment score in scientific subjects (mathematics, physics-chemistry, and biology); belong to the same quintile of average 9th grade continuous assessment

18. We do not look at the exact date of birth but only at whether the students have repeated at least one year: age is broken down to just one dummy variable.

score in humanities (French, history and foreign languages)¹⁹; and belong to the same decile of average 9th grade continuous assessment score across all subjects listed above.²⁰ Note that continuous assessment scores are binned into deciles or quintiles in order to have enough cases where at least two students share the exact same values for all observed characteristics. We show in section 2.5.1 that our estimates are unsensitive to using narrower or wider bins.

This definition of SF groups leaves us with 32,492 *groups* of SF students, 13,680 of which include students who were not assigned by their principal in the same 10th grade class.²¹ The total sample of SF students for which class effects can be estimated thus comprises 28,053 students out of an initial population of 2,897,986 10th grade students covering eight cohorts.²²

Model (2.1) is estimated on this "SF sample" of 28,053 students that have at least one similar-file mate ("SF mate") while ending up in separate classes, corresponding to 13,680 separate natural experiments. Each student is only compared to his SF mate since we control for α_g fixed effects. This single fixed effect controls altogether for years, class and middle school of origin, high school of destination and all other characteristics observed in students' registration files. *Therefore, there is no need to control separately for high school or year fixed effects, or for individual characteristics.*

Overall, these students are allocated to 22,162 different classes. Over the whole SF sample,

19. The foreign languages score is the weighted average of the student's main foreign language (weight = 2/3) and second foreign language (weight = 1/3). Using different weights does not change the paper's results. Note that the continuous assessment score in history is missing for 5.4 percent of observations. For these students, the average humanities score is the average of the French and foreign languages scores only.

20. Two students who belong to the same average score decile across all subjects may have very different subject-specific profiles: one may have high marks in sciences but not in humanities, or vice versa. Principals most probably differentiate between such students. This explains why we add quintiles of scientific and humanities scores separately, aside from the average score decile.

21. A total of 8,341 of the 32,492 SF groups are characterized by a set of optional courses that are only available in one high school class. Thus, these groups could not have been split up in any case. We conclude that the principals split up 13,680 SF groups out of 24,151 (57 percent) groups that could be split. The reason for separating out a group of SF students across different classes might be endogenous to potential outcomes, but does not affect our strategy. We only assume that, conditional on separating out SF students across different classes, principals decide randomly which students they assign to which class.

22. This population excludes 10th grade repeaters, but also newcomers for whom data on 9th grade exam scores is missing. Note that our definition of SF groups allows only 1 percent of the population to come to high school with at least one other student (only one out of 93 percent) who shares the same values for X_i , while ending up in different classes. This illustrates just how much our identification approach calls for a very rich database: only a large initial pool of students can yield a sample of SF students large enough to get accurate estimates of peer effects.

10,461 students are assigned to the same class as at least one student from another SF group. When estimating model (2.1), standard errors are therefore clusterized by 10th grade class, although this only impacts standard errors at the margin as the clusters are very small in size.

2.3.2 Empirical evidence of random assignment

The definition of SF groups we just described is already very restrictive. Considering principals' time constraint, they may not look for additional information to decide which one of the two students they assign to which class. As we described in section 2.2 however, there is still a few information they have that might be used. Assume, for instance, that there is substantial variation in ethnicity captured by names, even conditional on g , and that these variations are correlated with potential outcomes. If they are also taken into account by the principal when assigning these similar-file students to different classes, and if they are correlated to ϵ_{igc} , then assumption (2.2) is false and our results would be biased. The same argument holds for teachers' small written comments on students' behavior, that we do not observe.

2.3.2.1 Balancing test using anonymous exam scores

A first way to provide supporting evidence of the validity of assumption (2.2) is to exploit one information we do observe that principals do not: the students' anonymous scores A_i obtained in the national DNB exam just before entering high school. If conditional on g , principals do assign students based on some information we do not observe *and* that is correlated with potential outcomes, then some correlation between class characteristics C_{igc} and anonymous exam scores A_i should be observed, again, conditional on g . For instance, if 9th grade teachers feel a student is disruptive enough to warrant mentioning it in their written report, then they probably underscore his or her performance in class (as measured by continuous assessment scores). Therefore, disruptive students should display higher anonymous exam scores on average than their SF mate(s) with no behavioral issues, since SF students have very close continuous assessment scores by construction. This can be shown empirically taking our data on 2006–2011 cohorts, for which the behavior score NVS is available. A regression of A_i on NVS_i controlling for g (which includes continuous assessment scores) exhibits a negative correlation of -0.059 with a 0.020 standard error. It shows that when teachers report that a student's behavior is worse

than his or her SF mate, this student gets a higher score in the anonymous DNB exam, proving that they have been underscored in class.

A_i can thus be used to examine whether assumption (2.2) of a random assignment conditional on g is credible. We do so by estimating the following model:

$$C_{igc} = \alpha_g + \beta \cdot A_i + u_{igc} \quad (2.3)$$

where C_{igc} is any of class c 's characteristics and g is the index denoting "similar-file" (SF) groups, *i.e.* the groups of students sharing a specific vector of values for X_i . Adding α_g to the model constrains the regression to compare students solely with their SF mate(s). Under assumption (2.2), we expect β to be equal to zero.

TABLE 2.2 reports the estimated β parameters of model (2.3) for the entire SF sample for a number of class characteristics C_{igc} (columns I and II). Column I measures the raw sample correlations between ability and class characteristics, *i.e.* without the α_g fixed effect. In the SF sample, more able students are assigned to larger classes and with more persistent classmates; their classmates are also higher-achieving students, more often female and high-SES. All these correlations are statistically significant at the 1 percent level, except for the number of females. However, these correlations vanish within SF groups: as soon as we include the SF fixed effect, the estimates for β become very small and non-significant for all class characteristics (column II). In other words, for students who were similar with respect to X_i at the time of class assignment in a given high school and year, remaining differences in ability (unobserved by principals) have no correlation with differences between class characteristics.²³

This is a very strong result in favor of assumption (2.2). It clearly suggests that principals do not use any achievement-related information we do not observe to decide on class assignment for separated SF students, and thus split them up randomly (or at least exogenously to potential outcomes). In actual fact, this result is far from unrealistic since SF students are similar across

23. We provide two additional tests in the online appendix. First (TABLE 2.A.1), we show that the results hold when we use more detailed class characteristics regarding the number of persistent classmates of each type (low- or high-ability, same or opposite gender). Second (TABLE 2.A.2), we estimate equation (2.3) the other way around, *i.e.* regressing A_i on all class characteristics C_{igc} at the same time, thus measuring *partial correlations* between ability and each of the class characteristics. The conclusions of both these tests are identical to TABLE 2.2.

TABLE 2.2 – Student’s class characteristics regressed on own anonymous exam score: Evidence of the random assignment of similar-file students

	All	All	At risk	At risk
Dependent variable	(I)	(II)	(III)	(IV)
Number of PCs	0.062*** (0.003)	0.000 (0.005)	0.037*** (0.005)	−0.008 (0.008)
Average DNB score ¹	0.040*** (0.001)	0.001 (0.001)	0.047*** (0.001)	0.002 (0.002)
Number of girls	0.004 (0.006)	0.001 (0.011)	−0.014 (0.014)	−0.030 (0.020)
Number of high-SES students	0.316*** (0.008)	0.013 (0.009)	0.255*** (0.014)	0.019 (0.015)
Class size	0.067*** (0.004)	0.002 (0.006)	0.063*** (0.009)	0.000 (0.012)
<i>N</i>	28,053	28,053	8,981	8,981
SF fixed effect	No	Yes	No	Yes

Each cell is from a separate regression of the class characteristic of interest on the student’s standardized average anonymous score at the DNB exam.

The “at risk” sample consists of low-ability, low-SES students, on which the main effects’ magnitudes are the highest.

The SF fixed effect is a single fixed effect accounting altogether for 9th grade class and middle school of origin, high school of destination and individual characteristics, as defined in section 2.3.1. By controlling for the SF fixed effect, we compare each student only with her similar-file mate assigned randomly to another class.

All regressions include quadratic controls for the share of retained students and of missing DNB scores.

Standard errors (in parentheses) are clusterized at the 10th grade class level.

¹ This is the normalized DNB score. The normalization is done over the whole population: the sample’s mean is 0.245.

many variables, and the remaining vector of information is very small in comparison (names and teachers' written comments). It is absolutely consistent with our field observations that principals do not have time (or do not take time) to differentiate between very similar students.

Another possible scenario is that principals do consider the remaining information and that the latter are indeed correlated with potential high school outcomes, but *without being correlated with DNB test scores*. In other words, testing for imbalances in A_i would not be relevant since anonymous DNB scores are not a good measure of all unobserved determinants ϵ_{igc} of achievement in high school.²⁴ Yet we argue that this is very unlikely, since principals do not observe DNB scores. It is therefore hard to imagine that principals would use the few information we do not observe to separate SF students in a way that is correlated with ϵ_{igc} but without *any* correlation with A_i as shown in [TABLE 2.2](#).

[TABLE 2.2](#) also reports the results of the exogeneity test for a subsample of "at risk" SF students (columns III and IV). We define them as SF students who are low-achievers (below the median score of their middle school of origin) and low-SES. As we show in section [2.4.2.1](#), our main results regarding the positive effect of keeping classmates in the transition to high school are driven mostly by this specific subsample. For this reason, we check that the exogeneity test performs well for these students, as reported in columns III and IV of [TABLE 2.2](#). Therefore, we conclude that the SF students driving our main results are credibly exogenously assigned to their classes.

All in all, our definition of SF groups appears to be most suitable to estimate the causal effect of class environment on students' outcomes. Checking the robustness of our results to other specifications, we show in section [2.5.1](#) that our results barely change when using alternative specifications that are more or less restrictive.

24. For instance, a student's level of autonomy may be more important to high school achievement than achievement in the DNB exam.

2.3.2.2 Additional evidence of random assignment

When defining the g groups of SF students, we do not require students to share a similar behavioral score (*Note de vie scolaire* or NVS). This is because this score is not available for the first two cohorts, so including it would force us to remove one-quarter of our sample. However, **TABLE 2.2** suggests that it does not constitute a threat to our identification assumption. Otherwise, the resulting allocation of SF students would create a correlation between C_{igc} and A_i conditional on X_i , since anonymous DNB scores are correlated with behavioral issues conditional on teachers' grades (see above).

Additional evidence that principals do not take behavioral considerations to differentiate SF students can yet be provided. This is done by checking whether differences in class characteristics C_{igc} are correlated with potential differences in the NVS score in cohorts 2006 to 2011 (when the NVS score is available). We do so by estimating (2.3) after substituting A_i with a dummy for having an NVS score beneath the 10th percentile (equal to 15 out of 20).²⁵ Results are presented in **TABLE 2.3**. Most correlations between student behavior and class characteristics are both very small and non-significant at standard levels as long as comparisons are restricted within groups of SF students. This is true for both the entire SF sample (column II compared to column I) and the subsample of SF students at risk who drive our main results (column IV compared to column III).

Lastly, we run additional tests whose results are not reported for the sake of brevity. Firstly, all the analysis work presented in this paper is repeated for the 2006-2011 cohorts, constraining SF students to share a similar NVS score. Similar results are systematically found, though the estimates are less accurate. Secondly, we check whether differences in the continuous variables included in the definition of g groups (notably continuous assessment scores) conditional on their binned values are correlated with class characteristics C_{igc} . Once again, the correlations are mostly small and not significantly different from 0, especially for the subsample of students at risk of underachievement.

25. The NVS score has a very specific, negatively skewed distribution. 33 percent of students have the maximum score of 20 since they exhibit no disruptive behavior. The average score is 18, while the median score is 19. Therefore, we choose to define students with disruptive behavior as students with a score below the 10th percentile, which is precisely equal to 15 out of 20. Our results are not sensitive to the choice of threshold.

TABLE 2.3 – Student’s class characteristics regressed on behavior score: Evidence of the random assignment of similar-file students

	All	All	At risk	At risk
Dependent variable	(I)	(II)	(III)	(IV)
Number of PCs	−0.173*** (0.059)	−0.105 (0.084)	−0.117* (0.061)	0.017 (0.091)
Average DNB score	−0.157*** (0.014)	−0.005 (0.013)	−0.129*** (0.017)	−0.012 (0.018)
Number of girls	−0.557*** (0.136)	−0.039 (0.162)	−0.706*** (0.188)	0.115 (0.224)
Number of high-SES students	0.027 (0.202)	0.065 (0.122)	−0.023 (0.189)	−0.028 (0.165)
Class size	−0.280*** (0.093)	−0.030 (0.110)	−0.179 (0.116)	−0.023 (0.126)
<i>N</i>	15,775	15,775	4,952	4,952
SF fixed effect	No	Yes	No	Yes

Each cell is from a separate regression of the class characteristic of interest on a dummy for having an NVS score beneath the 10th percentile (equal to 15 out of 20).

The behavior score is not available for the first two cohorts, hence the smaller sample size.

The “at risk” sample consists of low-ability, low-SES students, on which the main effects’ magnitudes are the highest.

The SF fixed effect is a single fixed effect accounting altogether for 9th grade class and middle school of origin, high school of destination and individual characteristics, as defined in section 2.3.1. By controlling for the SF fixed effect, we compare each student only with her similar-file mate assigned randomly to another class.

All regressions include quadratic controls for the share of retained students and of missing DNB scores.

Standard errors (in parentheses) are clusterized at the 10th grade class level.

Overall, the empirical evidence strongly supports assumption (2.2) that high school principals randomly assign separated SF students (as defined in section 2.3.1) to their classes, or at least exogenously to achievement potential outcomes. The separation of similar-file students across 10th grade classes creates differences in educational outcomes that can therefore be attributed on average to differences in class characteristics.

2.3.3 Description of the SF sample

Descriptive statistics on the SF sample compared to the initial population are presented in column II of TABLE 2.1. All in all, students from the SF sample appear to be slightly higher achievers than the population of high school freshmen as a whole. Yet the differences are not always large in magnitude, although they are statistically significant. For example, the average DNB test score is 25.1 (sd = 5.6) in the SF sample compared to 23.9 (sd = 5.1) for the population as a whole. 15.0 percent of the SF sample repeat 10th grade as opposed to 15.3 percent of the total population, a difference that is again very small.²⁶ Column III reports the same descriptive statistics for the subsample of SF students at risk. By construction, these students are very low down on the ability distribution. They have an average normalized DNB score of -0.76 , repeat grade 10 almost 2.5 times as often as the average student in the population, and graduate from high school almost half as much.

2.4 Results

2.4.1 Freshman-year class characteristics and achievement

Given assumption (2.2), differences in class characteristics between SF mates are orthogonal to differences in individual unobservable characteristics. Therefore, conditional on g , regressing outcomes on any class characteristic – *e.g.* classmates' average ability – identifies a contextual

26. In terms of schools attended, SF students are found in 1,851 out of 2,679 high schools, *i.e.* 69 percent of all high schools. The high schools that do not get SF students are mostly very small schools, in which the chances of getting two students with the same registration files are small. They have 66 students on average in grade 10, versus 259 on average for high schools with SF students. Overall, the high schools containing SF students account for 91 percent of all 10th grade students.

effect that is not attributable to unobserved individual characteristics. Yet since classes differ in several ways simultaneously, the result of such a regression could be driven by some correlated, omitted class characteristics – *e.g.* the number of females. Hence as a first step, we attempt to figure out which aspect of the class environment is correlated with achievement, by regressing outcomes on several observed characteristics at once. We estimate model (2.1) where C_{igc} is a vector of peer characteristics commonly studied in the literature (average ability,²⁷ number of female students, number of high-SES students and class size), completed by the student's number of persistent classmates.

We consider different outcomes measured throughout the high school curriculum. The four first outcomes pertain to students' possible outcomes at the end of freshman year: repeating grade 10, dropping out,²⁸ enrolling in an academic or technological major.²⁹

Results are reported in columns I to IV of [TABLE 2.4](#). The number of persistent classmates is positively associated with achievement at the end of grade 10. On average, each additional persistent classmate a student gets with regard to her similar-file mate reduces her risk of repetition by -0.3 percentage point (pp.) and similarly raises enrollment in an academic major (se = 0.1 pp.), with small and non-significant effects on drop-out and enrollment in a technological major.³⁰ One additional female classmate reduces the risk of repetition by 0.1 pp. on average. This positive relationship between female peers and school achievement is consistent with the results found by other studies ([Hoxby, 2000](#); [Lavy and Schlosser, 2011](#)).³¹ Classmates' average

27. As measured by the DNB score. Because this data is missing for all repeating students (around 10 percent of classmates) and for another 20 percent of classmates (not matched, see section 2.2.3.1), we also include quadratic controls for the shares of repeating students and missing data.

28. As described in section 2.2, this "drop-out" measure picks up attrition due both to matching issues and actual drop-out. Since class environment is unlikely to substantially affect the matching procedure though, we believe this measure adequately captures the effect of class characteristics on the risk of drop-out.

29. We also estimated model (2.1) without controlling for the SF fixed effect to get the raw sample correlations. Basically, the number of persistent classmates displays positive, significant correlations with all outcomes, with larger estimates than those obtained with model (2.1). However, contrary to [TABLE 2.4](#)'s estimates, the classmates' average ability and female share are respectively positively and negatively associated with achievement. Detailed results on raw correlations are reported in the online appendix ([TABLE 2.A.3](#)).

30. We examine whether one academic major drives the effect, but find the same positive, non-significant effect on enrollment in sciences, humanities and social sciences. Results on specific major enrollment are not reported for the sake of brevity, but are available on request.

31. When adding an interaction term between own gender and the number of female classmates, we find that this effect is driven entirely by female students (no effect on males). Note that controlling for this interaction term does not change the estimate of the PC effect. This rules out the interpretation of the PC effect as capturing the

TABLE 2.4 – Effect of class characteristics on high school outcomes

	Repeats 10th grade	Drops out	Academic major	Tech. major	Takes <i>Bac</i> on schedule	HS graduate
Independent variable	(I)	(II)	(III)	(IV)	(V)	(VI)
Number of PCs	−0.003** (0.001)	−0.001 (0.001)	0.003** (0.001)	0.001 (0.001)	0.005** (0.002)	0.004* (0.002)
Average DNB score	0.040*** (0.010)	−0.012* (0.006)	−0.022** (0.010)	−0.006 (0.010)	−0.011 (0.014)	0.007 (0.014)
Number of girls	−0.001** (0.001)	0.001 (0.000)	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)	0.001 (0.001)
Number of high-SES students	−0.000 (0.001)	0.001** (0.001)	−0.000 (0.001)	−0.001 (0.001)	−0.000 (0.001)	0.001 (0.001)
Class size	−0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	−0.000 (0.001)	0.002 (0.002)	0.000 (0.002)
R^2	0.68	0.56	0.79	0.63	0.68	0.71
N	28,053	28,053	28,053	28,053	22,946 ¹	22,946 ¹
SF fixed effect	Yes	Yes	Yes	Yes	Yes	Yes

Each column is from a separate regression of students' outcomes on their class characteristics.

The SF fixed effect is a single fixed effect accounting altogether for 9th grade class and middle school of origin, high school of destination and individual characteristics, as defined in section 2.3.1. By controlling for the SF fixed effect, we compare each student only with her similar-file mate assigned randomly to another class.

All regressions include quadratic controls for the share of retained students and of missing DNB scores. Standard errors (in parentheses) are clusterized at the 10th grade class level.

¹ *Bac* data was not available for the last two cohorts, hence the smaller sample size.

ability is negatively associated with performance. A one standard deviation in the classmates' average DNB score³² significantly raises the risk of repetition by 3.8 pp. on average, but reduces both the drop-out rate and the probability of enrollment in an academic major. Peer ability displays effects that are therefore unclear.³³ The number of high-SES students also has a negative effect as it significantly increases the risk of dropping out, but its magnitude is rather small. Finally, we find no class size effect, most likely because of the small variance between classes in a given high school (the standard deviation of class size is only 1.9 students within SF groups).

TABLE 2.4 also reports results for two outcomes measured later than the end of 10th grade. Column V shows the effect of freshman-year class characteristics on the probability of students taking the *baccalauréat* exam "on schedule", *i.e.* three years after entering high school, meaning that they do not repeat grade 10 or grade 11 and that they make it through grade 12 without dropping out. Then, column VI investigates whether students with more persistent classmates in grade 10 are also more likely to pass the exam at that time. Interestingly, only the number of persistent classmates has a clear and enduring effect over time. Three years after entering high school, SF students who gain an additional persistent classmate in their freshman year are still more likely to take the *baccalauréat* exam at the end of grade 12. This result implies that the reduction in 10th grade repetition is not cancelled out by a higher propensity to repeat grade 11 or drop out of grade 12. Furthermore, they do not seem to perform any worse than others over these years, since they are also more likely to graduate from high school. In comparison, all other class characteristics display estimates that are rather small in magnitude and never statistically significant. Thus, the number of PCs seems highly relevant to capture the dimension(s) of class environment that matter, even more so than other, classic peer characteristics commonly studied in the literature.

Yet it is unclear what the number of PCs actually measures or captures. SF students' per-

impact of assignment to same-sex classmates.

32. The standard deviation of classmates' average ability *within SF groups* is only 27 percent of an average DNB score standard deviation.

33. The results obtained in the literature as regards the effect of peer ability are also mixed and inconclusive. Here, the negative peer effect is consistent with the impact of a lower relative position within the class, because students may look weaker to teachers when assigned with better classmates. This may have little effect on drop-outs, but it would raise their risk of repeating grade 10 and at the same time reduce their chances of admission on an academic major track.

sistent classmates could affect them by means of mechanisms implying all sorts of unobserved characteristics that generate peer effects, such as ability and motivation. In the following, we provide strong evidence suggesting that the PC effect does not capture an ability peer effect, but rather works through a social network mechanism. As we will show, the most consistent interpretation of the data is that students simply benefit from getting peers they know and with whom they are used to interacting.

2.4.2 The protective role of familiarity with classmates

We first check that controlling or not for other class characteristics does not affect the PC estimates. In [TABLE 2.5](#), we report the previous estimates of the effect of PCs from regressions where C_{igc} is the full vector (column II). In column I, the effect of the number of PCs is estimated without controlling for the other class characteristics. The estimates are virtually identical for the two regressions, indicating that PCs are not correlated with these other class characteristics.³⁴ This first piece of evidence strongly suggests that the number of persistent classmates does not capture any other omitted dimension of class environment. Suppose, for instance, that SF students' persistent classmates had specific characteristics associated with higher performance, which benefit SF students without any link with familiarity. Considering [TABLE 2.5](#)'s results, such characteristics should not be correlated with DNB score, gender or SES. In other words, to be consistent with [TABLE 2.5](#)'s pattern, any credible omitted class characteristic driving the PC effect would have to be uncorrelated with all other observed class dimensions in C_{igc} . It is highly unlikely that such a characteristic exists.³⁵ In section 2.5.2, we run a robustness check that takes variations in PC within-classes to check that the PC estimate is unlikely to be driven by characteristics that are fixed at class level.

In the remaining tables of the present section, we systematically include quadratic controls for other class characteristics when we estimate the effect of PCs. As seen in [TABLE 2.5](#), these controls do not affect our estimates.

34. This remains true even if we allow the model to account for non-linearities in the effect of peer ability, *e.g.* by controlling for the share of very high- (or low-) achieving peers.

35. Unfortunately, we do not have access to teachers' characteristics at class level. Yet any correlation between PCs and teachers' characteristics should drive a change in the PC estimate when controlling for students' average ability, because assignment of teachers to classes is mostly related to students' ability.

TABLE 2.5 – Effect of persistent classmates on high school outcomes with and without controlling for other class characteristics

Dependent variable	(I)	(II)
Repeats 10th grade	−0.003** (0.001)	−0.003*** (0.001)
Drops out	−0.001 (0.001)	−0.001 (0.001)
Academic major	0.003** (0.001)	0.003** (0.001)
Technological major	0.001 (0.001)	0.001 (0.001)
<i>N</i>	28,053	28,053
Takes <i>Bac</i> on schedule	0.005*** (0.002)	0.005*** (0.002)
HS graduate	0.004** (0.002)	0.004* (0.002)
<i>N</i> ¹	22,946	22,946
SF fixed effect	Yes	Yes
Control for other class characteristics	No	Yes

Each cell is from a separate regression of students' outcomes on their number of PCs, controlling or not for the other class characteristics presented in TABLE 2.4. Column (II) is identical to the first row in TABLE 2.4.

The SF fixed effect is a single fixed effect accounting altogether for 9th grade class and middle school of origin, high school of destination and individual characteristics, as defined in section 2.3.1. By controlling for the SF fixed effect, we compare each student only with her similar-file mate assigned randomly to another class.

All regressions include quadratic controls for the share of retained students and of missing DNB scores.

Standard errors (in parentheses) are clusterized at the 10th grade class level.

¹ *Bac* data was not available for the last two cohorts, hence the smaller sample size.

2.4.2.1 Distribution of the PC effect

Another reason to believe in the "familiarity" interpretation concerns the heterogeneity of the PC effect. If historical familiarity with classmates really matters, it is probably not equally important for all students. In particular, we expect that the role of former classmates is all the more important for those students who are likely to experience a difficult transition. This is exactly what we find.

We provide an analysis of the distribution of the PC effect in [TABLE 2.6](#). We first investigate how it varies with SF students' level of achievement (panel B). To retain enough statistical power, we split the sample into just two parts, defining SF students as either low- or high-ability based on their relative position with respect to the median continuous assessment score for their middle school of origin.³⁶ The PC effect is strikingly heterogeneous across these two categories. While the number of PCs has virtually no effect on high-ability students, low-ability students are strongly and positively impacted. For the sake of brevity, we do not report on the magnitude of the estimates, since the effects are actually highly heterogeneous again between low- and high-SES within this subgroup. As reported in panel C, the effects observed on low-ability students are almost exclusively driven by low-SES students. On average, each additional PC reduces their risk of repetition by 1.4 pp, though not their risk of dropping out. They are therefore significantly more likely to enroll for either an academic or a technological major, with a similar increase in magnitude. No backlash to this strong short-term impact can be found in following grades. On average, each PC in freshman year raises their chances of taking the *baccalauréat* exam and graduating by the same amount. In comparison, the estimates are very small in magnitude and never statistically significant at conventional levels for low-ability high-SES students (column V) and high-ability low-SES students (not reported). This suggests that keeping some classmates matters only for students who may be experiencing a hard transition *both* academically – they were already performing poorly in middle school – and culturally – their parents come from the working class and might not have studied at high school.

36. Therefore, the terms "low-ability" and "high-ability" do not denote distribution ends. Besides, we use the continuous assessment score since SF groups are defined with regard to it, such that two SF students are necessarily both below or above the median. Although the anonymous DNB exam score would be a better measure of ability, two SF students may be on different sides of the median DNB score. We would thus lose part of the SF sample by analyzing the PC effect separately on each side of the median DNB score. However, doing so brings us to the same conclusions as in [TABLE 2.6](#), though the estimates are often less accurate.

TABLE 2.6 – Distribution of the PC effect

Dependent variable	Repeats 10th grade	Drops out	Academic major	Tech. major	Takes <i>Bac</i> on schedule	HS graduate
Independent variable	(I)	(II)	(III)	(IV)	(V)	(VI)
(A) Whole sample: average effect						
Number of PCs	−0.003*** (0.001)	−0.001 (0.001)	0.003** (0.001)	0.001 (0.001)	0.005*** (0.002)	0.004* (0.002)
<i>N</i>	28,053	28,053	28,053	28,053	22,946 ¹	22,946 ¹
(B) Whole sample: by ability						
Number of PCs	−0.000 (0.001)	−0.001 (0.001)	0.002 (0.001)	−0.001 (0.001)	0.002 (0.002)	0.001 (0.002)
Low-ability × PC	−0.009*** (0.003)	−0.001 (0.002)	0.003 (0.003)	0.006** (0.003)	0.009** (0.004)	0.007 (0.004)
<i>N</i>	28,053	28,053	28,053	28,053	22,946 ¹	22,946 ¹
(C) Low-ability students: by SES						
Number of PCs	0.002 (0.005)	−0.002 (0.003)	−0.003 (0.005)	0.003 (0.003)	0.003 (0.007)	0.000 (0.007)
Low-SES × PC	−0.015** (0.006)	0.001 (0.004)	0.011* (0.007)	0.003 (0.005)	0.011 (0.009)	0.012 (0.008)
<i>N</i>	11,383	11,383	11,383	11,383	9,588 ¹	9,588 ¹
(D) Low-ability, low-SES students ("at risk"): by disruption intensity						
Number of PCs	−0.010* (0.006)	0.003 (0.004)	0.006 (0.005)	0.001 (0.005)	0.007 (0.006)	0.010* (0.006)
Disruptive transition × PC	−0.009 (0.008)	−0.009 (0.006)	0.006 (0.008)	0.012 (0.007)	0.015 (0.010)	0.005 (0.009)
<i>N</i>	8,981	8,981	8,981	8,981	7,615 ¹	7,615 ¹
(E) Low-ability, low-SES students ("at risk"): by gender						
Number of PCs	−0.018*** (0.006)	0.000 (0.005)	0.014** (0.006)	0.004 (0.005)	0.010 (0.008)	0.009 (0.008)
Girl × PC	0.007 (0.009)	−0.002 (0.006)	−0.009 (0.008)	0.004 (0.007)	0.007 (0.010)	0.005 (0.010)
<i>N</i>	8,981	8,981	8,981	8,981	7,615 ¹	7,615 ¹

Each column in each panel is from a separate regression.

The SF fixed effect is a single fixed effect accounting altogether for 9th grade class and middle school of origin, high school of destination and individual characteristics, as defined in section 2.3.1. By controlling for the SF fixed effect, we compare each student only with her similar-file mate assigned randomly to another class.

All regressions include quadratic controls for the share of retained students, the share of missing DNB scores, and the class characteristics presented in TABLE 2.4.

Standard errors (in parentheses) are clusterized at the 10th grade class level.

¹ *Bac* data was not available for the last two cohorts, hence the smaller sample size.

In the rest of the paper, we derive results only for this specific subsample of low-ability, low-SES students, which we call students "at risk". We focus on this subsample because the effect of persistent classmates is entirely driven by these students. As in TABLE 2.5, the presence or not of other class characteristics on the right hand side of the regression does not change the estimates of the PC effect.

In panel D, we investigate this distribution pattern further by looking into how the PC effect varies with the difference in school-level social environment (measured by the share of high-SES students). This gap, denoted Δp , is negative for one-third of low-ability low-SES students only.³⁷ These low-ability, low-SES students are twice as likely to experience a positive Δp , meaning that they enter an environment with more high-SES students than they used to have.

Although the sample size is too small to draw definite conclusions (the differences between the two groups are not significant to the 10 percent level), we do observe a difference in magnitude between students at risk depending on the sign of Δp . The difference goes in the direction we would expect, *i.e.* that students who experience a more difficult transition ($\Delta p > 0$) are more sensitive to the presence of persistent classmates.

In panel E, we also estimate the difference of the PC effect between male and female students at risk. However, the results do not display any clear heterogeneity in the gender dimension. Persistent classmates seem to have more impact on the male repetition rate than the female repetition rate. Yet the discrepancy runs in the opposite direction for the *baccalauréat* outcomes, with larger estimates for females. Both male and female students thus seem to benefit from persistent classmates in freshman year, although the benefits differ slightly depending on the stage of the high school curriculum.³⁸

37. This is a mechanical consequence of the lower probability of finding low-SES students enrolled in general high schools after grade 9.

38. Although the increase in academic major enrollment is similar across both genders, persistent classmates switch male outcomes from repetition to science majors only and female outcomes to humanities only. To be more precise, both male and female PCs increase male enrollment in science, while females enroll more in humanities only when they get more female PCs. Results available on request.

We further analyzed the distribution of the effect in middle and high school contexts. These results are not reported since no other interesting pattern can be found. For example, the effect does not appear to vary significantly with middle or high school size, the share of middle school classmates attending the high school, or the 10th grade class context.³⁹

All in all, the results of this investigation are consistent with our interpretation of the PC effect. The estimates reported in [TABLE 2.4](#) are a watered-down version of the very strong PC effect on SF students who experience an upheaval in the transition to high school.⁴⁰ Already knowing some peers in the class matters a lot to low-ability students with low socioeconomic status who come from an environment that is poor compared to the high school. This is most consistent with the interpretation of the PC effect as reflecting the impact of familiarity. By contrast, it is unlikely that the former classmates of these low-achieving underprivileged students have higher unobservables than average, which would drive the PC effect. The following section presents additional evidence in support of this.

2.4.2.2 Do all former peers matter?

If the effect of persistent classmates is explained by familiarity, then students at risk should be more affected by peers with whom they have been more likely to interact at middle school.⁴¹ In particular, they may have interacted much more with their former classmates than with middle school peers in other classes. In [TABLE 2.7](#), panel A, we add to the previous regressions the number of these former *middle school mates from other classes*. We find a small, negative effect on grade repetition, but it is not significant. Surprisingly, this effect is related to a small increase in the risk of dropping out, statistically significant at the 5 percent level. Other estimates are very

39. We check, in particular, whether the extent to which your new classmates are grouped with their former classmates increases your need to be with yours. Yet again, we find no result to suggest that this is the case. This is noteworthy as it suggests that grouping former classmates would not drive negative spillovers on their other classmates, who do not necessarily have many former classmates in high school. It would be helpful, however, to confirm such a conclusion with a controlled field experiment to directly investigate externalities within the class.

40. We check whether the other peer characteristics studied in [TABLE 2.4](#) also have a larger effect on this specific category of students "at risk". Results are provided in the online appendix ([TABLE 2.A.4](#)). Again, other peer characteristics (average ability, number of females, etc.) display non-significant and non-persisting effects on high school achievements, even for these students.

41. We focus solely on the subsample of SF students at risk since section [2.4.2.1](#) shows that they are the only ones driving the PC effect. The balancing test shown in [TABLE 2.2](#) is repeated for this subsample for the different types of PCs in the online appendix ([TABLE 2.A.1](#)). The test is satisfied for all types of PCs.

small in magnitude and never statistically significant. Therefore, students seem to benefit only from their middle school mates who were in the same class, with whom they probably interacted much more.

Students do not appear to benefit more from some specific types of persistent classmates. After controlling for the number of all persistent classmates, the number of same-gender PC (panel B) and high-ability PC (panel C) do not trigger different effects on academic achievement. Interestingly, high-ability PCs do not appear to be more beneficial to students. Therefore, the PC effect is highly unlikely to result from a higher unobserved ability of persistent classmates.⁴² This is another evidence that students benefit from their persistent classmates only because they know each other.

2.5 Robustness checks

2.5.1 Alternative SF group specifications

All the results presented in section 2.4 are based on a quite restrictive definition of SF groups. We required students to have the exact same values for all variables that were included in the registration files, binning test scores into deciles or quintiles. This degree of accuracy was necessary in order for our exogeneity tests to be valid. In this section however, we explore the sensitivity of the results for alternative definitions of the SF groups.

We test both definitions that are more restrictive and less restrictive than the main definition. For instance, we can be less restrictive by allowing students to have different values for some of the variables observed by principals or by broadening the bins of the continuous test scores. Conversely, we can be more restrictive by using narrower bins. In TABLE 2.8, we report the effect of persistent classmates on grade repetition for four alternative specifications for the primarily affected sample of "at risk" students (low-ability, low-SES). The reference definition

42. We again define "high-ability" as students with continuous assessment scores above the school median, for consistency with TABLE 2.6. However, anonymous exam scores are a better measure of ability and could be used here to define classmates' ability without any loss of accuracy (by comparison to SF students, see again footnote 36). The same results are found by using one or another measure.

TABLE 2.7 – Which peers do matter? Decomposition of the PC effect on students at risk.

Dependant variable	Repeats 10th grade	Drops out	Academic major	Tech. major	Takes <i>Bac</i> on schedule	HS graduate
Independent variable	(I)	(II)	(III)	(IV)	(V)	(VI)
(A) Persistent classmates and persistent schoolmates						
Number of PCs	−0.014*** (0.004)	−0.001 (0.003)	0.008* (0.004)	0.006* (0.004)	0.013*** (0.005)	0.011** (0.005)
Persistent schoolmates from other classes	0.000 (0.003)	0.002 (0.002)	−0.001 (0.002)	−0.001 (0.002)	−0.004 (0.003)	−0.003 (0.003)
(B) Persistent classmates by gender						
Number of PCs	−0.015** (0.007)	−0.001 (0.004)	0.007 (0.007)	0.008 (0.006)	0.014* (0.007)	0.011 (0.007)
Same sex PCs	0.001 (0.010)	−0.001 (0.006)	0.002 (0.009)	−0.002 (0.009)	−0.001 (0.011)	0.003 (0.010)
(C) Persistent classmates by ability						
Number of PCs	−0.017** (0.007)	−0.001 (0.004)	0.014** (0.006)	0.005 (0.006)	0.014* (0.008)	0.011 (0.007)
High-ability PCs	0.007 (0.010)	0.000 (0.006)	−0.010 (0.009)	0.003 (0.009)	−0.000 (0.011)	0.001 (0.011)
<i>N</i>	8,981	8,981	8,981	8,981	7,615 ¹	7,615 ¹

The sample is limited to low-ability, low-SES students ("at risk") sample on which the effects' magnitudes are the highest.

Each cell is from a separate regression of students' outcomes on their number of persistent classmates.

The SF fixed effect is a single fixed effect accounting altogether for 9th grade class and middle school of origin, high school of destination and individual characteristics, as defined in section 2.3.1. By controlling for the SF fixed effect, we compare each student only with her similar-file mate assigned randomly to another class.

All regressions include quadratic controls for the share of retained students, the share of missing DNB scores, and the class characteristics presented in TABLE 2.4.

Standard errors (in parentheses) are clusterized at the 10th grade class level.

¹ *Bac* data was not available for the last two cohorts, hence the smaller sample size.

is reproduced in column IV; columns I to III show the results for less restrictive definitions and column V presents a more restrictive definition. The details of each definition are given in the table.

All specifications yield a significant negative effect of the number of persistent classmates on grade repetition, although the magnitude of the effect varies from one specification to another. Bear in mind that the balancing test presented in section 2.3 produces less conclusive results for the alternative specifications.⁴³ The results in column IV therefore remain our reference results. However, the fact that the effect retains the same sign and size is reassuring for the validity and robustness of our results.

2.5.2 Estimation based on the impact of SF student allocation on their classmates

In section 2.4, the identification of peer effects within classes is based directly on the comparison of SF students randomly assigned to different classes. Yet the random assignment of SF students may also be seen itself as a shock to class composition. Receiving one or another SF classmate in the class can make a difference to the other students in these classes. A focus on the effects of SF student allocation on their 10th grade classmates yields new estimates of the effect of classmate persistence, hence testing the robustness of the results presented in section 2.4.

2.5.2.1 Principle

Using the notations from **FIGURE 2.2**, we now compare students C to H with each other instead of comparing student A to student B. If A and B are defined as similar-file, the result of the random allocation of A and B should have no impact on students C–H, since A and B have the same characteristics. However, if we allow A and B to come from *two different classes* of the same middle school, the result of the allocation will have an impact on students C to H if some of them come from A or B's 9th grade class. Given the similarity of A and B from most points of view except their specific class of origin, the result of the allocation will only affect this

43. Results available on request.

TABLE 2.8 – Robustness check: effect of PC on low-ability students’ repetition rate using different specifications of the SF fixed effect

Independent variable	Specifications				
	(I)	(II)	(III)	(IV)	(V)
PC	−0.004*** (0.001)	−0.012*** (0.004)	−0.015*** (0.004)	−0.014*** (0.004)	−0.019*** (0.007)
R^2	0.58	0.68	0.60	0.61	0.62
N	169,258	19,369	11,404	8,981	3,214
SF students share...					
Options	✓	✓	✓	✓	✓
Middle school	✓	✓	✓	✓	✓
9th grade class	Indifferent	Same	Similar	Same	Same
In-school score	Decile	Decile	Decile	Decile	Decile
Science score		Quintile	Quintile	Quintile	Decile
Humanities score		Quintile	Quintile	Quintile	Decile
Held back	✓	✓	✓	✓	✓
Gender			✓	✓	✓
2-category SES			✓	✓	✓

The sample is limited to low-ability, low-SES students (“at risk”) sample on which the effects’ magnitudes are the highest.

Each cell is from a separate regression of students’ outcomes on their number of persistent classmates.

All regressions include quadratic controls for the share of retained students, the share of missing DNB scores, and the class characteristics presented in [TABLE 2.4](#). Standard errors (in parentheses) are clusterized at the 10th grade class level.

All regressions include similar-file fixed effects, which are different in each column. Column IV is the original definition of SF groups, columns I to III are less restrictive and column V is more restrictive.

dimension of the class characteristics vector C_{ic} . For example, if students A and C come from the same 9th grade class, C would have one more PC in case 1 than in case 2.

Therefore, in this section, we use another definition of SF students in which we require all variables to be identical *except for the classroom of origin* (the middle school or origin must be the same). However, we find in our exogeneity test (2.3) that principals do distinguish between students who come from different classes if those classes do not have similar characteristics. Therefore, in this section, students will be considered similar-file students if they come from classes that share three characteristics: quintile of the average DNB score, proposing or not elite options, and quintile of the number of students going to the destination high school.

Formally for each SF group j produced by the new definition, we can define an instrument Z_{ij} equal to the number of persistent classmates that student i obtains from that particular SF group. The variable is defined only for students who were in the same 9th grade class as one of the SF students in group j , and who are in one of the 10th grade classes attended by these SF students. We denote this sample $\mathcal{P}_j$. Note that in this context, i belongs to the sample of *former classmates* in the SF sample, $\mathcal{P} = \bigcup_j \mathcal{P}_j$, and not to the SF sample itself.

Formally, we estimate the following reduced form model:

$$Y_{ijkc} = \alpha_{jk} + \beta \cdot Z_{ij} + \epsilon_{ijkc} \quad (2.4)$$

where c denotes the 10th grade class and k denotes the 9th grade class. The α_{jk} fixed effect ensures that comparisons are made between students who belong to the same $\mathcal{P}_j$ sample and come from the same 9th grade class. Alternatively, this fixed effect can be replaced with a α_{jc} fixed effect, where we compare solely students ending up in the same 10th grade class. β identifies the causal effect of getting one additional persistent classmate.⁴⁴

2.5.2.2 Validity of the test

The exogeneity of this instrument is based on a stronger assumption than the main model. Although SF students are randomly split, the assignment of other freshmen is not exogenous.

44. Z_{ij} is a "perfect" instrument for PC, as it has a correlation of one with PC and there is no compliance issue here. This is why we estimate the reduced form model directly.

In particular, if they were assigned with respect to the number of PCs, instrument Z_{ij} would not be exogenous. Therefore, model (2.4) is identified under the hypothesis that $\mathcal{P}$ students are not allocated to classes correlated with PCs.

In order to check this additional hypothesis, we estimate the correlation between the value of instrument Z_{ij} (the number of persistent classmates received by random allocation) and the individual characteristics of the students $i \in \mathcal{P}_j$:

$$Z_{ij} = \alpha_{jk} + \gamma \cdot X_i + u_{ijkc} \quad (2.5)$$

where X_i is the vector of observable characteristics tested.

We present the results of this test in [TABLE 2.9](#). In column I, we find that individual characteristics are correlated with the instrument when the controls for α_{jk} fixed effects are not included. However, these correlations vanish when we include them (column II) or when we replace them with a 10th-grade-class fixed effect (column III).⁴⁵ These results suggest that the students who obtain a randomly assigned PC are comparable in their observed dimensions to those who do not, within $\mathcal{P}_j$ samples. Although these students might be different on an unobserved level, we argue that this test is satisfactory enough to run a robustness check on our main results presented in section 2.4.

2.5.2.3 Results

The results of the estimation of model (2.4) are produced in [TABLE 2.10](#). As in [TABLE 2.9](#), the α_{jk} fixed effect is omitted in column I, included in column II and replaced with a 10th-grade-class fixed effect in column III. Since Z_{kc} takes the same value for all students in 10th grade class c from the same 9th grade class k , the standard errors are clustered within kc groups. Similar to section 2.4, we find that a higher number of persistent classmates is associated with lower grade repetition and higher enrollment for the academic major. We also observe a long-term positive effect on high school graduation. The orders of magnitude are similar to the results using the first strategy and do not vary drastically depending on the fixed effect we include.

45. We use DNB quintile dummies instead of the DNB score to avoid losing students with missing values. Students for which the DNB score is missing have all five dummies equal to zero.

TABLE 2.9 – IV exogeneity test

Independent variable	(I)	(II)	(III)	(IV)
Held back	0.002 (0.017)	−0.005 (0.016)	0.007 (0.015)	−0.007 (0.015)
Girl	−0.006 (0.011)	0.004 (0.009)	0.011 (0.008)	−0.001 (0.008)
High-SES	0.064*** (0.011)	0.009 (0.009)	0.006 (0.008)	0.013* (0.008)
High quality optional course	0.044* (0.023)	0.003 (0.022)	0.012 (0.018)	0.004 (0.018)
DNB Quintile 1	Ref. —	Ref. —	Ref. —	Ref. —
DNB Quintile 2	0.011 (0.017)	−0.022 (0.016)	−0.006 (0.015)	−0.009 (0.016)
DNB Quintile 3	0.042** (0.018)	−0.017 (0.016)	−0.000 (0.016)	−0.013 (0.017)
DNB Quintile 4	0.082*** (0.019)	0.005 (0.017)	0.016 (0.017)	0.009 (0.017)
DNB Quintile 5	0.088*** (0.020)	0.009 (0.018)	0.009 (0.017)	0.014 (0.017)
R^2	0.01	0.23	0.44	0.43
N	28,515	28,515	28,515	28,515
Fixed effect	None	High school	HS × 9th grade class	10th grade class

Each column is from a separate regression of the instrument Z (number of PCs received through one of the random allocations) on students' characteristics.

Standard errors (in parentheses) are clusterized at the class of origin $\times$ class of destination level.

TABLE 2.10 – Effect of PC on high school outcomes using the IV strategy

Dependent variable	(I)	(II)	(III)	(IV)
Repeats 10th grade	−0.014*** (0.003)	−0.009*** (0.003)	−0.007** (0.003)	−0.008*** (0.003)
R^2	0.00	0.18	0.32	0.32
Drops out	0.000 (0.002)	−0.001 (0.002)	0.001 (0.002)	−0.003 (0.002)
R^2	0.00	0.12	0.25	0.25
Academic major	0.029*** (0.004)	0.012*** (0.004)	0.008** (0.004)	0.012*** (0.004)
Technological major	−0.016*** (0.002)	−0.002 (0.002)	−0.002 (0.002)	−0.001 (0.002)
N	33,663	33,663	33,663	33,663
Takes <i>Bac</i> on schedule	0.013*** (0.005)	0.012*** (0.004)	0.007 (0.005)	0.014*** (0.005)
HS graduate	0.028*** (0.005)	0.016*** (0.005)	0.011** (0.005)	0.016*** (0.005)
N	28,515	28,515	28,515	28,515
Fixed effect	None	High- school	HS × 9th grade class	10th grade class

Each column is from a separate regression of students' outcomes on the instrument Z (number of PCs received through one of the random allocations).

Standard errors (in parentheses) are clusterized at the class of origin $\times$ class of destination level.

All in all, these results confirm our main strategy results. Moreover, this approach has some advantages even though it relies on a stronger assumption. First, the main strategy focuses on students who have been separated from a very similar former classmate, most likely a friend. Getting more persistent classmates may have more of an impact than usual in such settings. This approach finds a similar impact on a different sample, thus removing doubts as to the external validity of our results. Furthermore, by allowing comparisons of students within the same class (column III), the effect can be estimated of a pure variation in the number of PCs, holding other class characteristics constant.⁴⁶ This allays our concerns about omitted class characteristics driving our results in section 2.4. Last but not least, it shows that the positive effect of persistent classmates does not operate solely through improvement in the global class context, which would affect everyone similarly.⁴⁷ Freshmen do therefore benefit from familiar peers via channels that operate at individual level, such as a greater sense of belonging or social and academic support.

2.6 Discussion and conclusion

This paper documents how classes influence students' achievements in high school. Empirical evidence suggests that freshmen students with very similar registration files, when separated among different classes, are randomly assigned to their class. With this quasi-experimental setting, differences in class environments can be credibly assumed orthogonal to potential outcomes. After examining the correlations between a number of measures of class composition and student outcomes, we find a robust and significant effect of being assigned again with more former classmates. Yet this effect is all but homogenous. It is almost exclusively driven by low-achieving, low-SES freshmen who enroll in high schools with more high-SES mates than they used to have. These students "at risk" in high school do not benefit more from the presence of high-achieving persistent classmates. It may be a surprising result, since these peers are more likely to provide academic help compared to low-achieving persistent classmates. Most probably, low-achieving students are better-off by former classmates through social channels.

46. Suppose in [FIGURE 2.2](#) that C comes from A's class and D from B's class. For C and D, getting A or B only changes their relative number of PCs. This is true *by construction*, because A and B have the same characteristics as regards all other dimensions.

47. It might have been expected, for example, that it is more comfortable to teach a class if more students already know each other in the class. This beneficial impact on teachers might then affect all students in the class, even those who are not directly affected by having more former classmates. Yet in this case, we would not find any difference between students in the same 10th grade class.

Mechanisms implying direct interactions could be at work. For example, persistent classmates could be friends or acquaintances to whom freshmen may talk during the early weeks, ask for help, or even work as a team.⁴⁸ However, even where there is no interaction, being surrounded by peers they know and who experience the same difficulties may also be a psychological relief, fostering their sense of belonging in the high school. Precise data on students' relationships and well-being such as those used by [Lavy and Sand \(2012\)](#) would help understand how freshmen take advantage of familiarity with peers. Unfortunately, such data is not available in our context.

Most importantly, though, our results show that students do not bear the brunt of increased enrollment in grade 11 with lower performance in subsequent years. So, whatever the mechanisms at work, we know that persistent classmates do raise achievement.⁴⁹ Basically, we argue that this result suffices to draw relevant policy recommendations on class composition. Whereas a great deal of money is usually invested in improving outcomes of students at risk of underachievement, we show that assigning them to some persistent classmates could increase their performance at no cost. The potential gains could be substantial. Our analysis finds that each persistent classmate reduces their risk of repetition by 1.5 pp. on average. This figure is estimated using solely the variance of PCs observed within groups of SF students at risk, with 98 percent of variations ranging from 0 to 3 PCs (no conclusion should be drawn about the PC effect beyond this range). Students at risk in the freshmen population have 1.5 low-achieving PCs on average: raising this figure by 3⁵⁰ could thus reduce their risk of repetition by 4.5 pp. (meaning 13 percent of their current rate) while increasing their graduation rate by the same amount. Non-linearities of the PC effect might result in the same (or even a greater) benefit with fewer than 3 PCs, but the small variance in the number of PCs in our sample does not allow us to investigate this. Nonetheless, it would be helpful to back up these recommendations with experimental evidence, as [Carrell et al. \(2011\)](#) stressed the need to test such policy predictions

48. In this way, grouping students who come from the same class may be an efficient tool to help teachers develop cooperative learning within the class, as they might rely on existing friendships and social links between students right from the start of the year.

49. In particular, mechanisms implying solely a change in preferences are ruled out by this result. For example, persistent classmates could make students less likely to repeat 10th grade only by increasing their propensity to appeal to teachers' decisions (to avoid losing friends). In that case though, negative drawbacks would be observed in the following grades since students would enroll in grades and for majors where requirements would be too high.

50. 11 former classmates are enrolled in their high school on average.

drawn from reduced-form estimates.

We believe this study makes an important contribution to the existing literature on the role of the school environment in achievement. We show that low-achieving students benefit from peer persistency when the rest of their environment gets largely disrupted by the transition to high school. This study emphasizes the need for a minimum of stability in the face of great instability, and highlights the huge complexity of peer effects and social interactions.

2.7 Bibliography

- Angrist, J. D. and Lang, K. (2004). Does school integration generate peer effects? evidence from boston's metco program. *The American Economic Review*, 94(5):1613–1634.
- Calvò-Armengol, A., Patacchini, E., and Zenou, Y. (2009). Peer Effects and Social Networks in Education. *Review of Economic Studies*, 76(4):1239–1267.
- Carrell, S. E., Sacerdote, B. I., and West, J. E. (2011). From natural variation to optimal policy? the lucas critique meets peer effects. NBER Working Paper 16865, National Bureau of Economic Research.
- Cullen, J. B., Jacob, B. A., and Levitt, S. (2006). The effect of school choice on participants: Evidence from randomized lotteries. *Econometrica*, 74(5):1191–1230.
- De Giorgi, G. and Pellizzari, M. (2013). Understanding Social Interactions: Evidence from the Classroom. NBER Working Papers 19202, National Bureau of Economic Research, Inc.
- Duflo, E., Dupas, P., and Kremer, M. (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in kenya. *The American Economic Review*, 101(5):1739–74.
- Foster, G. (2006). 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(8–9):1455 – 1475.

- Fruehwirth, J. C. (2013). Identifying peer achievement spillovers: Implications for desegregation and the achievement gap. *Quantitative Economics*, 4(1):85–124.
- Gibbons, S., Silva, O., and Weinhardt, F. (2013). I (don't) like the way you move: The disruptive effects of residential turnover on student attainment. Preliminary draft, SOLE 2014 Conference.
- Goux, D. and Maurin, É. (2007). Close neighbours matter: Neighbourhood effects on early performance at school. *Economic Journal*, 117(523):1192–1215.
- Halliday, T. J. and Kwak, S. (2012). What is a peer? the role of network definitions in estimation of endogenous peer effects. *Applied Economics*, 44(3):289–302.
- Hoxby, C. M. (2000). Peer effects in the classroom: Learning from gender and race variation. NBER Working Paper 7867, National Bureau of Economic Research.
- Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika*, 87(3):706–710.
- Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.
- Lavy, V., Paserman, M. D., and Schlosser, A. (2012a). Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *Economic Journal*, 122(559):208–237.
- Lavy, V. and Sand, E. (2012). The friends factor: How students' social networks affect their academic achievement and well-being? NBER Working Paper 18430, National Bureau of Economic Research.
- Lavy, V. and Schlosser, A. (2011). Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 3(2):1–33.
- Lavy, V., Silva, O., and Weinhardt, F. (2012b). The good, the bad, and the average: Evidence on ability peer effects in schools. *Journal of Labor Economics*, 30(2):367–414.
- Mora, T. and Oreopoulos, P. (2011). Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends. *Economics of Education Review*, 30(4):575–581.

2.A Additional tables

TABLE 2.A.1 – Student’s class characteristics regressed on own anonymous exam score: Evidence of the random assignment of similar-file students

	All	All	At risk	At risk
Dependent variable	(I)	(II)	(III)	(IV)
Number of PCs	0.062*** (0.003)	0.000 (0.005)	0.037*** (0.005)	−0.008 (0.008)
Same sex PCs	0.037*** (0.002)	0.004 (0.004)	0.019*** (0.003)	−0.002 (0.006)
Opposite sex PCs	0.025*** (0.002)	−0.003 (0.003)	0.018*** (0.003)	−0.006 (0.005)
High-ability PCs	0.050*** (0.002)	0.005 (0.004)	0.029*** (0.003)	−0.005 (0.006)
Low-ability PCs	0.012*** (0.001)	−0.004 (0.003)	0.007** (0.003)	−0.003 (0.005)
<i>N</i>	28,053	28,053	8,981	8,981
SF fixed effect	No	Yes	No	Yes

Each cell is from a separate regression of the number of PCs of each type on the student’s standardized average anonymous score at the DNB exam.

The “at risk” sample consists of low-ability, low-SES students, on which the main effects’ magnitudes are the highest.

The SF fixed effect is a single fixed effect accounting altogether for 9th grade class and middle school of origin, high school of destination and individual characteristics, as defined in section 2.3.1. By controlling for the SF fixed effect, we compare each student only with her similar-file mate assigned randomly to another class.

All regressions include quadratic controls for the share of retained students and of missing DNB scores.

Standard errors (in parentheses) are clusterized at the 10th grade class level.

TABLE 2.A.2 – Anonymous exam score regressed on class characteristics: Evidence of the random assignment of similar-file students

	All	All	At risk	At risk
Independent variable	(I)	(II)	(III)	(IV)
Number of PCs	0.149*** (0.013)	0.000 (0.012)	0.059*** (0.022)	−0.025 (0.026)
Average DNB score	4.572*** (0.069)	0.088 (0.084)	3.065*** (0.092)	0.094 (0.141)
Number of girls	−0.022*** (0.006)	0.000 (0.007)	−0.010 (0.008)	−0.018 (0.012)
Number of high-SES students	−0.010* (0.005)	0.008 (0.008)	0.024*** (0.009)	0.015 (0.016)
Class size	0.046*** (0.011)	−0.001 (0.012)	−0.008 (0.013)	0.005 (0.020)
R^2	0.25	0.90	0.17	0.78
F -test	0.00	0.47	0.00	0.35
N	28,053	28,053	8,981	8,981
SF fixed effect	No	Yes	No	Yes

Each column is from a separate regression of the the student's standardized average anonymous score at the DNB exam on class characteristics.

The "at risk" sample consists of low-ability, low-SES students, on which the main effects' magnitudes are the highest.

The SF fixed effect is a single fixed effect accounting altogether for 9th grade class and middle school of origin, high school of destination and individual characteristics, as defined in section 2.3.1. By controlling for the SF fixed effect, we compare each student only with her similar-file mate assigned randomly to another class.

All regressions include quadratic controls for the share of retained students and of missing DNB scores.

Standard errors (in parentheses) are clusterized at the 10th grade class level.

TABLE 2.A.3 – Raw correlation between class characteristics and high school outcomes

Dependant variable	Repeats 10th grade	Drops out	Academic major	Tech. major	Takes <i>Bac</i> on schedule	HS graduate
Independent variable	(I)	(II)	(III)	(IV)	(V)	(VI)
Number of PCs	−0.006*** (0.001)	−0.001 (0.000)	0.013*** (0.001)	−0.007*** (0.001)	0.008*** (0.001)	0.011*** (0.001)
Average DNB score	−0.072*** (0.005)	−0.025*** (0.003)	0.176*** (0.006)	−0.079*** (0.005)	0.108*** (0.007)	0.177*** (0.007)
Number of girls	0.001*** (0.000)	−0.001*** (0.000)	0.000 (0.001)	−0.000 (0.000)	0.000 (0.001)	0.000 (0.001)
Number of high-SES students	0.001*** (0.000)	0.000** (0.000)	0.001** (0.000)	−0.003*** (0.000)	−0.002*** (0.001)	0.001 (0.001)
Class size	−0.003*** (0.001)	0.000 (0.000)	0.003*** (0.001)	−0.000 (0.001)	0.005*** (0.001)	0.004*** (0.001)
R^2	0.03	0.01	0.09	0.04	0.04	0.08
N	28,053	28,053	28,053	28,053	22,946 ¹	22,946 ¹
SF fixed effect	No	No	No	No	No	No

Each column is from a separate regression of students' outcomes on their class characteristics.

The SF fixed effect is not included in these regressions.

All regressions include quadratic controls for the share of retained students and of missing DNB scores.

Standard errors (in parentheses) are clusterized at the 10th grade class level.

¹ *Bac* data was not available for the last two cohorts, hence the smaller sample size.

TABLE 2.A.4 – Effect of class characteristics on high school outcomes for the sample "at risk"

	Repeats 10th grade	Drops out	Academic major	Tech. major	Takes <i>Bac</i> on schedule	HS graduate
Independent variable	(I)	(II)	(III)	(IV)	(V)	(VI)
Number of PCs	−0.014*** (0.004)	−0.001 (0.003)	0.009** (0.004)	0.007* (0.004)	0.014*** (0.005)	0.012** (0.005)
Average DNB score	0.064** (0.025)	−0.004 (0.014)	−0.053** (0.021)	−0.007 (0.023)	−0.037 (0.028)	−0.018 (0.026)
Number of girls	−0.004* (0.002)	0.001 (0.001)	0.001 (0.002)	0.002 (0.002)	0.002 (0.002)	0.001 (0.002)
Number of high-SES students	−0.000 (0.003)	0.001 (0.002)	−0.000 (0.002)	−0.001 (0.002)	−0.001 (0.003)	0.004 (0.003)
Class size	0.000 (0.003)	0.001 (0.002)	0.001 (0.003)	−0.003 (0.003)	0.001 (0.004)	−0.001 (0.004)
R^2	0.61	0.56	0.67	0.58	0.61	0.60
N	8,981	8,981	8,981	8,981	7,615 ¹	7,615 ¹
SF fixed effect	Yes	Yes	Yes	Yes	Yes	Yes

The sample is limited to low-ability, low-SES students ("at risk") sample on which the effects' magnitudes are the highest.

Each column is from a separate regression of students' outcomes on their class characteristics.

The SF fixed effect is a single fixed effect accounting altogether for 9th grade class and middle school of origin, high school of destination and individual characteristics, as defined in section 2.3.1. By controlling for the SF fixed effect, we compare each student only with her similar-file mate assigned randomly to another class.

All regressions include quadratic controls for the share of retained students and of missing DNB scores. Standard errors (in parentheses) are clusterized at the 10th grade class level.

¹ *Bac* data was not available for the last two cohorts, hence the smaller sample size.

Chapitre 3

La ségrégation sociale et scolaire en France : un état des lieux quantitatif

Ce chapitre a été écrit avec Son Thierry Ly.

Il sera intégré au rapport du Conseil national d'évaluation du système scolaire sur l'évolution des inégalités scolaires.

Nous remercions Cédric Afsa, Pauline Givord, Marine Guillerm, Arthur Heim, Catherine Moisan, Nathalie Mons, Olivier Monso et Fabrice Murat pour leurs nombreuses suggestions qui ont permis d'enrichir ce rapport.

3.1 Introduction

La mixité sociale et scolaire est un objectif affirmé par de nombreux acteurs du système éducatif, que ce soit au niveau politique ou des acteurs de terrain. Cet objectif est inscrit dans la loi

de refondation de l'école et il a été réitéré par la ministre de l'Éducation nationale début 2015. Une enquête du Cnesco montre par ailleurs que la très grande majorité des chefs d'établissements souhaitent éviter de constituer des classes homogènes au sein des établissements. Pourtant, les chiffres montrent l'existence d'une véritable ségrégation sociale et scolaire dans le système éducatif français. En effet, les collégiens et lycéens d'origine aisée comptent en moyenne dans leur classe deux fois plus de camarades également d'origine aisée que les autres élèves. Alors qu'une classe de 25 élèves contient en moyenne 5 élèves « CSP+ »¹, un élève lui-même classé CSP+ en comptera en moyenne 10 dans sa classe. 20 % des élèves « favorisés » de troisième comptent même plus de 15 élèves favorisés dans leur classe, alors que seuls 2 % des élèves des classes moyennes et défavorisées en comptent autant².

Ces chiffres sont le résultat d'un décalage entre des objectifs collectifs de mixité et une recherche individuelle d'entre-soi. À l'échelle individuelle en effet, on comprend que les parents d'élèves aisés cherchent pour leur enfant un environnement scolaire rassurant. De même, un enseignant peut estimer qu'il est plus agréable, sinon plus efficace, de faire cours à une classe où les écarts de niveau sont limités. Un chef d'établissement convaincu que des classes homogènes sont inefficaces doit donc faire face à des demandes qui viennent s'opposer à l'objectif de mixité.

L'effet de l'environnement de classe ou d'établissement sur les performances scolaires est très difficile à évaluer, et il n'y a pas de consensus au niveau de la communauté scientifique sur cette question³. Il en est de même pour les acteurs de terrain dont les convictions varient fortement en fonction des expériences vécues par chacun. La recherche de mixité scolaire répond cependant à d'autres impératifs : l'école doit proposer les mêmes opportunités à tous les élèves, elle doit les préparer à devenir des citoyens et à vivre ensemble.

1. On appelle « CSP+ » les élèves de la catégorie « Favorisés-A » définie dans le **TABLE 3.1.**

2. Ces chiffres portent sur la cohorte entrée en sixième en 2007 et sont calculés à partir des bases administratives du Ministère de l'Éducation nationale, qui sont la principale source de cette étude.

3. Très souvent, des élèves ayant des entourages différents ont un contexte scolaire qui diffère par beaucoup d'aspects. Dès lors, les comparaisons de résultats d'élèves ne permettent pas d'isoler l'effet de la composition de classe ou d'établissement. Les études sur les effets de pair présentent souvent des défauts méthodologiques ou une validité limitée à des cas très précis, et leurs résultats ne peuvent pas être généralisés. Dans ce rapport, nous ne cherchons pas à identifier l'impact de la mixité sociale et scolaire sur les résultats scolaires mais seulement à la quantifier et l'analyser.

Si la mixité sociale est scolaire est devenue une préoccupation importante au cours des dernières années, l'Éducation nationale ne s'est pas dotée d'un appareil statistique permettant de la mesurer précisément. La Depp⁴ et l'Insee ont commencé à proposer une méthodologie (Monso et al., 2015) mais sans intention, à ce stade, de produire des indicateurs de manière régulière. L'objectif de ce rapport est de dresser un panorama complet de la ségrégation sociale et scolaire dans l'enseignement secondaire français, afin d'éclairer les débats actuels sur la mixité sociale à l'école⁵. Il vient combler un manque dans la littérature en économie et sociologie de l'éducation, en étant le premier à proposer un diagnostic :

- portant sur l'ensemble du territoire national ;
- sur le collège et le lycée ;
- mesurant la ségrégation sociale et scolaire ;
- intégrant les dimensions inter- et intra-établissement.

Plusieurs études ont cependant permis d'introduire des premières données chiffrées dans le débat. Les premiers éléments quantitatifs précis ont été apportés par Felouzis (2003), qui a analysé la répartition des 144 000 élèves scolarisés dans les 333 collèges de l'académie de Bordeaux pendant l'année scolaire 2000-2001. À partir de la base élèves académique, il a défini une variable ethnique en utilisant le prénom des élèves, en distinguant élèves « autochtones » et « allochtones » (prénoms d'origine étrangère, 7,1 % des élèves). Cette variable a permis de mesurer le degré de regroupement de ces élèves dans certains collèges : par exemple, 10 % des collèges scolarisent 26 % des élèves allochtones. Pour atteindre l'équilibre parfait, il faudrait en théorie demander à 4,9 % de l'ensemble des élèves de l'académie de changer d'établissement, soit plus de 7 000 élèves – l'équivalent de seize collèges. Malheureusement, l'analyse de Felouzis s'arrête au niveau établissement ; pour atteindre un équilibre entre toutes les *classes* de l'académie, il faudrait ensuite déplacer les élèves entre les classes de chaque établissement.

Citons par ailleurs plusieurs travaux réalisés à la suite de l'assouplissement de la carte scolaire en 2007. Thauvel-Richard and Murat (2013) ont montré une hausse des demandes de déroga-

4. Direction de l'évaluation, de la prospective et de la performance, le service statistique du Ministère de l'Éducation nationale.

5. Une version préliminaire de ce rapport a été diffusée lors de la conférence de comparaison internationale du Cnesco sur les mixités sociales à l'école, les 4 et 5 juin 2015. Les principaux ajouts de cette version concernent le rôle du privé dans la ségrégation inter-établissement, une analyse actualisée et plus précise du rôle des options dans la ségrégation intra-établissement, et des données sur l'évolution de la ségrégation au cours de la dernière décennie.

tion autour de la date de l'assouplissement de la carte scolaire et des stratégies d'évitement des collèges de l'éducation prioritaire qui bénéficie notamment aux collèges privés. Ce phénomène, bien qu'ayant augmenté en amplitude, n'a pas eu un effet détectable sur les chiffres macroscopiques de la ségrégation au niveau national (Fack and Grenet, 2012). Il a toutefois pu aggraver les disparités entre les *collèges* au niveau local, comme à Paris (Merle, 2010) ou à Lille et Saint-Étienne (Ben Ayed et al., 2013). À Paris, l'introduction en 2008 du logiciel d'affectation des élèves Affelnet a engendré dans les lycées parisiens une baisse de la ségrégation *sociale*, en raison d'un bonus pour les boursiers dans le barème Affelnet de cette académie, mais le poids des notes dans ce barème a au contraire augmenté la ségrégation *scolaire* (Fack et al., 2014).

Au niveau intra-établissement, les études quantitatives sont très rares, en raison notamment du manque de données. Un des premiers résultats provient d'une analyse sur un échantillon de 212 collèges entre 1989 et 1992 (Duru-Bellat and Mingat, 1997), qui met en avant une diversité de pratiques au sein des établissements, donnant lieu à la constitution de classes de niveau dans un collège sur deux environ. Nous avons montré dans une précédente étude portant sur la région Île-de-France (Ly et al., 2014) qu'en moyenne, un élève de milieu favorisé compte dans sa classe deux fois plus d'élèves eux-mêmes de milieu favorisé qu'un élève de milieu moyen ou défavorisé n'en compte – un résultat que nous confirmons ici au niveau national. Dans cette région, la ségrégation intra-établissement est du même ordre de grandeur que la ségrégation entre les établissements d'une même commune, mais celle-ci s'explique en grande partie par le « hasard » de la constitution des classes : en simulant une affectation aléatoire des élèves, on obtient des valeurs indicateurs de ségrégation intra-établissement proches des valeurs observées.

Nous commençons par présenter les données mobilisées et la méthodologie utilisée, en discutant notamment le choix d'un indicateur de ségrégation (partie 3.2). Nous présentons ensuite les mesures de ségrégation sociale et scolaire aux niveaux inter-établissements (partie 3.3) et intra-établissement (partie 3.4), puis l'évolution des niveaux de ségrégation (partie 3.5).

3.2 Méthodologie

3.2.1 Comment mesurer la ségrégation ?

La ségrégation, qu'elle soit sociale ou scolaire, est un concept difficile à définir et à mesurer. Dans une acception historique, ce terme désigne une politique volontariste de séparation des individus en fonction de leurs caractéristiques personnelles, par exemple leur couleur de peau. Pour les économistes et les sociologues, ce terme peut avoir un sens moins extrême et capter tout un continuum de situations « plus ou moins ségréguées ». Nous utiliserons la définition suivante :

Définition 1 (Ségrégation) *Le terme de ségrégation désigne toute situation dans laquelle des individus ayant des caractéristiques différentes fréquentent des environnements différents.*

Une telle situation se traduit d'une part par des disparités de répartition de groupes sociaux selon les lieux, d'autre part par une diminution des interactions entre individus issus de groupes différents (Massey and Denton, 1988).

Cette définition permet de définir des *indicateurs de ségrégation* dont le but est de mesurer l'intensité de ces disparités.

Définition 2 (Indicateur de ségrégation) *Un indicateur de ségrégation est un nombre compris entre zéro et un (ou entre 0 % et 100 %), dont la valeur minimale est atteinte lorsque les individus fréquentent un environnement semblable quelles que soient leurs caractéristiques personnelles et la valeur maximale est atteinte lorsqu'ils sont entièrement isolés en fonction de ces caractéristiques.*

Plusieurs indices de ségrégation respectant cette définition ont été construits pour étudier l'inégale répartition des individus dans différents contextes. Ils ont des fondements théoriques et mathématiques différents et donc des significations et des interprétations différentes (Frankel and Volij, 2011), sur lesquelles nous reviendrons succinctement. Ils ont cependant en commun de demander à ceux qui les utilisent d'effectuer deux « découpages » de la population :

- Un découpage en fonction des *caractéristiques personnelles* : il s'agit de définir des groupes sociaux correspondant au type d'isolement que l'on souhaite mesurer. Par exemple, pour

mesurer la ségrégation ethnique, on peut définir des groupes en fonction de l'origine migratoire ; pour mesurer la ségrégation sociale, on peut définir des groupes en fonction des revenus ou de la catégorie socio-professionnelle ; pour mesurer la ségrégation scolaire, on peut définir des groupes en fonction des notes obtenues à un examen. Dans tous les cas, les groupes doivent être tels que chaque individu dans la population étudiée doit appartenir à un groupe et un seul. Souvent, on choisit un découpage simple en *deux catégories uniquement*, pour faciliter la lecture des indices⁶.

- Un découpage en fonction des *environnements fréquentés* : il s'agit de définir l'environnement social de chaque individu, et l'échelle à laquelle on mesure la ségrégation. Dans le cas de l'école, on peut s'intéresser notamment à l'environnement établissement ou à l'environnement classe. Les indices de ségrégation obtenus en utilisant l'environnement classe seront toujours supérieurs aux indices obtenus en utilisant l'environnement établissement⁷. En effet, les disparités entre classes sont le résultat de deux phénomènes qui s'ajoutent : la ségrégation *inter-établissements*, qui est celle que l'on mesure en utilisant les environnements établissements, et la ségrégation *intra-établissement*, liée à la manière de composer les classes.

Dans cette étude, nous utiliserons quatre découpages en fonction des caractéristiques personnelles, qui serviront à définir deux indices de ségrégation sociale et deux indices de ségrégation scolaire :

Définition 3 (Ségrégation sociale) *La ségrégation sociale est mesurée en découplant la population en fonction de leur origine sociale. Nous effectuons deux tels découpages, qui permettent de mesurer l'isolement des élèves dits « CSP+ » d'une part, celui des élèves « CSP- » d'autre part.*

- Dans le premier découpage, le groupe A correspond aux élèves dits « CSP+ », c'est-à-dire dont le premier parent inscrit (souvent le père) a une catégorie socio-professionnelle classée dans le groupe « Favorisés-A » de la Depp et le groupe B correspond aux élèves dont le parent est classé dans un des groupes « Favorisés-B », « Moyens » et « Défavorisés » (voir **TABLE 3.1**)⁸. Ce dé-

6. L'étude de **Monso et al. (2015)** mesure un indice de ségrégation s'appuyant sur les quatre catégories sociales de la Depp.

7. Les deux indices peuvent en théorie être égaux dans le cas, jamais observé en pratique, d'une ségrégation absolument nulle au niveau intra-établissement.

8. Cette classification a été remise en question par une étude récente de la Depp (**Le Donné and Rocher, 2010**) qui montre que cette classification pourrait être revue, notamment en ce qui concerne les PCS défavorisées.

coupage permet de mesurer à quel point les élèves « CSP+ » sont isolés par rapport aux autres élèves.

- *Dans le deuxième découpage, le groupe A correspond aux élèves dits « CSP- », c'est-à-dire dont le parent de référence appartient à la catégorie « Défavorisés » de la Depp et le groupe B correspond aux élèves dont le parent est classé dans un des groupes « Favorisés-A », « Favorisés-B » et « Moyens ». Ce découpage permet de mesurer à quel point les élèves « CSP- » sont isolés par rapport aux autres élèves.*

Dans une situation hypothétique où les enfants de parents CSP+ sont fortement regroupés et que les autres élèves sont relativement mélangés (c'est-à-dire qu'il y a des établissements très favorisés et des établissements moyens), l'indice de ségrégation sociale « CSP+ » sera élevé alors que l'indice de ségrégation sociale « CSP- » sera modéré. À l'inverse, dans une situation où les enfants de parents CSP- sont fortement regroupés et que les autres élèves sont relativement mélangés (c'est-à-dire qu'il y a des établissements très défavorisés et des établissements moyens), l'indice de ségrégation sociale « CSP- » sera élevé alors que l'indice de ségrégation sociale « CSP+ » sera modéré.

Définition 4 (Ségrégation scolaire) *La ségrégation scolaire est mesurée en découplant la population en fonction de leur niveau scolaire. Cette dimension de la ségrégation est plus difficile à mesurer, dans la mesure où nous ne disposons que de mesures imprécises du niveau scolaire, à savoir les résultats au diplôme national du brevet (DNB) et au baccalauréat. Les résultats au baccalauréat sont difficiles à comparer entre les voies générale, technologique et professionnelle et entre les différentes séries au sein de ces voies. Une autre mesure, là encore imprécise, du niveau scolaire, est l'âge de l'élève, dans la mesure où les élèves les plus en difficulté sont susceptibles de redoubler au cours de leur scolarité. Comme pour la ségrégation sociale, nous effectuons deux tels découpages, qui permettent de mesurer l'isolement des « meilleurs élèves » d'une part, celui des élèves plus en difficulté d'autre part.*

- *Dans le premier découpage, le groupe A correspond aux élèves qui n'ont pas redoublé en sixième, cinquième et quatrième, ont passé le DNB dans la filière générale, et dont la note moyenne aux quatre épreuves finales (mathématiques, français, histoire-géographie, histoire de l'art) est dans le premier quart de la distribution ; le groupe B correspond à tous les autres élèves : ceux qui ont redoublé au collège, ceux qui passent le DNB dans la filière professionnelle, et ceux dont*

Code	Libellé	Catégorie
23	Chefs d'entreprise de 10 salariés ou plus	Favorisés-A
31	Professions libérales	
32	Cadres de la fonction publique, professions intellectuelles et artistiques	
36	Cadres d'entreprise	
42	Professeurs des écoles, instituteurs et assimilés	
43	Professions intermédiaires de la santé et du travail social	Favorisés-B
44	Clergé, religieux	
45	Professions intermédiaires administratives de la fonction publique	
46	Professions intermédiaires administratives et commerciales des entreprises	
47	Techniciens	
48	Contremaîtres, agents de maîtrise	
73	Anciens cadres et professions intermédiaires	Moyens
10	Agriculteurs	
21	Artisans	
22	Commerçants et assimilés	
51	Employés de la fonction publique	
54	Employés administratifs d'entreprise	
55	Employés de commerce	
56	Personnels des services directs aux particuliers	
71	Anciens agriculteurs exploitants	
72	Anciens artisans, commerçants, chefs d'entreprise	Défavorisés
61	Ouvriers qualifiés	
66	Ouvriers non qualifiés	
69	Ouvriers agricoles	
76	Anciens employés et ouvriers	
81	Chômeurs n'ayant jamais travaillé	
82	Inactifs divers (autres que retraités)	

TABLE 3.1 – Classification des PCS de la DEPP

les résultats aux épreuves finales ne se situent pas dans le premier quart de la distribution. Ce découpage permet de mesurer à quel point ces élèves, que nous appellerons « meilleurs élèves », sont isolés par rapport aux autres élèves. Attention, le niveau scolaire étant mesuré à la fin de la troisième, les chiffres présentés ici sont d'autant moins précis qu'ils concernent un niveau éloigné de la troisième : d'une part, les élèves peuvent avoir un niveau scolaire variable au cours de leur scolarité, et d'autre part nos données ne nous permettent pas toujours de suivre les élèves pendant l'ensemble de leur scolarité.

- *Il est plus compliqué de détecter les élèves les plus en difficulté en utilisant les résultats au DNB ; ainsi, dans le deuxième découpage, le groupe A correspond aux élèves ayant redoublé au moins une fois au cours de leur scolarité au moment de la mesure, c'est-à-dire aux élèves « en retard scolaire » ; le groupe B correspond donc aux élèves n'ayant jamais redoublé au moment de la mesure (élèves « à l'heure » ou en avance scolaire). Ce découpage permet de mesurer à quel point les élèves en retard scolaire sont isolés par rapport aux autres élèves.*

Notons que le brevet des collèges intervenant à la fin du collège, il faut rester prudent quant à l'interprétation des chiffres de la ségrégation scolaire au collège. Les inégalités de niveau scolaire mesurées entre établissements et entre classes peuvent être à la fois le résultat d'une allocation inégale des élèves à la rentrée et du fait que différentes classes ont progressé inégalement au cours de l'année, pour des raisons multiples (allocation des moyens, des professeurs, etc.). Les données dont nous disposons ne nous permettent malheureusement pas de dissocier ces deux composantes. Cela ne retire pas d'intérêt au fait de mesurer les inégalités scolaires au collège, qu'elles soient issues d'une véritable ségrégation au moment de l'allocation des élèves ou de dynamiques de classe différentes en cours d'année. Comme nous le verrons notamment dans le cas de la ségrégation intra-établissement, nous disposons par ailleurs d'indices nous indiquant que l'allocation initiale des élèves aux classes contribue de manière importante à la ségrégation scolaire. De plus, la mesure de la ségrégation scolaire basée sur la répartition des élèves en retard scolaire – qui fournit des résultats très similaires – reflète uniquement des inégalités de niveau scolaire au moment de la rentrée.

Dans cette étude, nous présenterons plus en détail les mesures de ségrégation des élèves « favorisés », c'est-à-dire des CSP+ et des meilleurs élèves. Sauf mention contraire, nous parle-

rons donc de ségrégation sociale pour désigner la ségrégation des élèves CSP+ et de ségrégation scolaire pour désigner la ségrégation des meilleurs élèves. Les résultats portant sur les élèves CSP- et les élèves en retard scolaire seront présentés en annexe.

3.2.2 L'indice d'exposition

La littérature en économie et sociologie utilise plusieurs indices de ségrégation : nous rappelons la définition des quatre plus populaires d'entre eux dans l'annexe 3.A. Nous avons choisi de travailler avec l'indice d'exposition (voir définition 5), pour deux raisons principales.

Tout d'abord, il peut s'écrire comme la somme d'une composante inter-établissements et d'une composante intra-établissements. Il permet donc de déterminer dans quelle mesure la ségrégation s'explique par les disparités entre établissements et par la composition des classes. Seul cet indice et l'indice de Theil (définition 7) possèdent cette propriété de décomposabilité.

Le deuxième intérêt de cet indice est qu'il possède une interprétation simple en termes d'environnements vécus par les élèves. Par exemple, si l'indice d'exposition mesure une ségrégation sociale entre les établissements de 20 %, cela signifie que le pourcentage d'élèves CSP+ dans le collège d'un élève lui-même CSP+ est supérieur de 20 points, en moyenne, que celui d'un élève non-CSP+. Ainsi, s'il y a 30 % d'élèves CSP+ dans la population et que la ségrégation sociale est de 20 %, les élèves CSP+ auront en moyenne 44 % d'élèves CSP+ dans leur établissement, alors que les autres élèves n'en auront que 24 %. L'indice d'exposition est donc une mesure directe de l'« entre-soi » scolaire : il indique dans quelle mesure les élèves fréquentent des élèves qui possèdent les mêmes caractéristiques qu'eux.

La possibilité de décomposer l'inter- et l'intra-établissement, l'interprétation intuitive de l'indice et d'autres propriétés mathématiques (pour plus de détails, voir [Ly and Riegert, 2015](#)) sont les raisons pour lesquelles nous privilégions l'utilisation de ce dernier indice dans cette étude ; à des fins de comparaison, les valeurs des autres indices sont donnés en annexe.

Par ailleurs, nous accompagnerons systématiquement la valeur de l'indice de ségrégation de la proportion du groupe de référence (CSP+ ou « meilleurs élèves ») dans l'échantillon étudié. Il est

important de noter qu'il n'existe pas de manière absolue de comparer deux indices de ségrégation mesurés sur des échantillons ayant des compositions sociales ou scolaires différentes, et ce quel que soit l'indice choisi (voir à ce sujet la discussion de l'annexe 3.A.2).

3.2.3 Les données mobilisées

Les résultats présentés dans cette étude sont issus de données de la Direction de l'évaluation, de la prospective et de la performance (Depp) du Ministère de l'Éducation nationale. Ils proviennent en particulier de deux bases de données :

- La base *Scolarité* recense l'ensemble des élèves inscrits dans les collèges et lycées français publics ou privés sous contrat. Pour chaque élève, sont connus son établissement d'inscription, sa classe, sa division (par exemple, « troisième A »), les langues et options qu'il a choisies, sa date de naissance, son sexe ainsi que la catégorie socio-professionnelle de ses parents (voir TABLE 3.1).
- La base *Océan-Brevet* contient les notes obtenues par l'ensemble des élèves de troisième au diplôme national du brevet (DNB).

Nous n'avons conservé dans l'échantillon que les classes du cursus scolaire classique, et avons notamment exclu les classes SEGPA⁹ et les autres classes spécialisées. Le but de cette sélection est de ne mesurer que la ségrégation qui a lieu de manière naturelle entre des classes *a priori* comparables, en mettant de côté les dispositifs d'adaptation et de remédiation qui ont pour effet direct de réduire la mixité sociale.

Les résultats présentés ici portent sur la pseudo-cohorte d'élèves qui entrent en sixième en 2007. On mesure donc la ségrégation en sixième sur l'année scolaire 2007-2008, la ségrégation en cinquième sur l'année 2008-2009, jusqu'à la ségrégation en terminale sur l'année 2013-2014.

9. Section d'enseignement général et professionnel adapté.

3.3 La ségrégation sociale et scolaire entre les établissements

3.3.1 Au niveau national

La **FIGURE 3.1** donne le niveau de ségrégation sociale et scolaire, niveau par niveau, au collège et au lycée, sur le plan national. Les points reliés indiquent, pour chaque niveau, la part d'élèves du « groupe de référence » (les élèves CSP+ pour la ségrégation sociale, les meilleurs élèves pour la ségrégation scolaire), et les barres indiquent le niveau de ségrégation entre les établissements. La partie plus claire de chaque barre indique la part de la ségrégation expliquée par les disparités entre communes.

Un collège compte en moyenne 110 élèves par niveau ; un lycée (tous types confondus) compte environ 192 élèves de seconde (ou première année de CAP), 172 élèves de première (ou deuxième année de CAP) et 162 élèves de Terminale. Selon nos critères, la part de CSP+ varie de 21 à 26 % et la part des meilleurs élèves est comprise entre 19 et 24 %, selon le niveau. On parlera dans la suite de *cohortes* pour désigner l'ensemble des élèves d'un établissement à un niveau donné (par exemple, l'ensemble des élèves de quatrième du collège Louise-Michel constitue une cohorte). L'exposition d'un élève au groupe de référence (les CSP+ pour la ségrégation sociale, les meilleurs élèves pour la ségrégation scolaire) est donc définie comme la part de ce groupe dans la cohorte de l'élève.

3.3.1.1 Ségrégation sociale

L'indice de ségrégation sociale au collège est compris entre 16 et 17 % selon le niveau. Alors qu'en l'absence de ségrégation, tous les élèves compteraient 22 % d'élèves CSP+ parmi les élèves du même niveau de leur collège, les élèves eux-même CSP+ en comptent 34 %, et les élèves des CSP intermédiaires et défavorisées n'en comptent que 18 %.

Au lycée, les chiffres de la ségrégation sociale n'augmentent que légèrement. Cette relative stabilité est le résultat de deux phénomènes concurrents : le regroupement dans les mêmes éta-

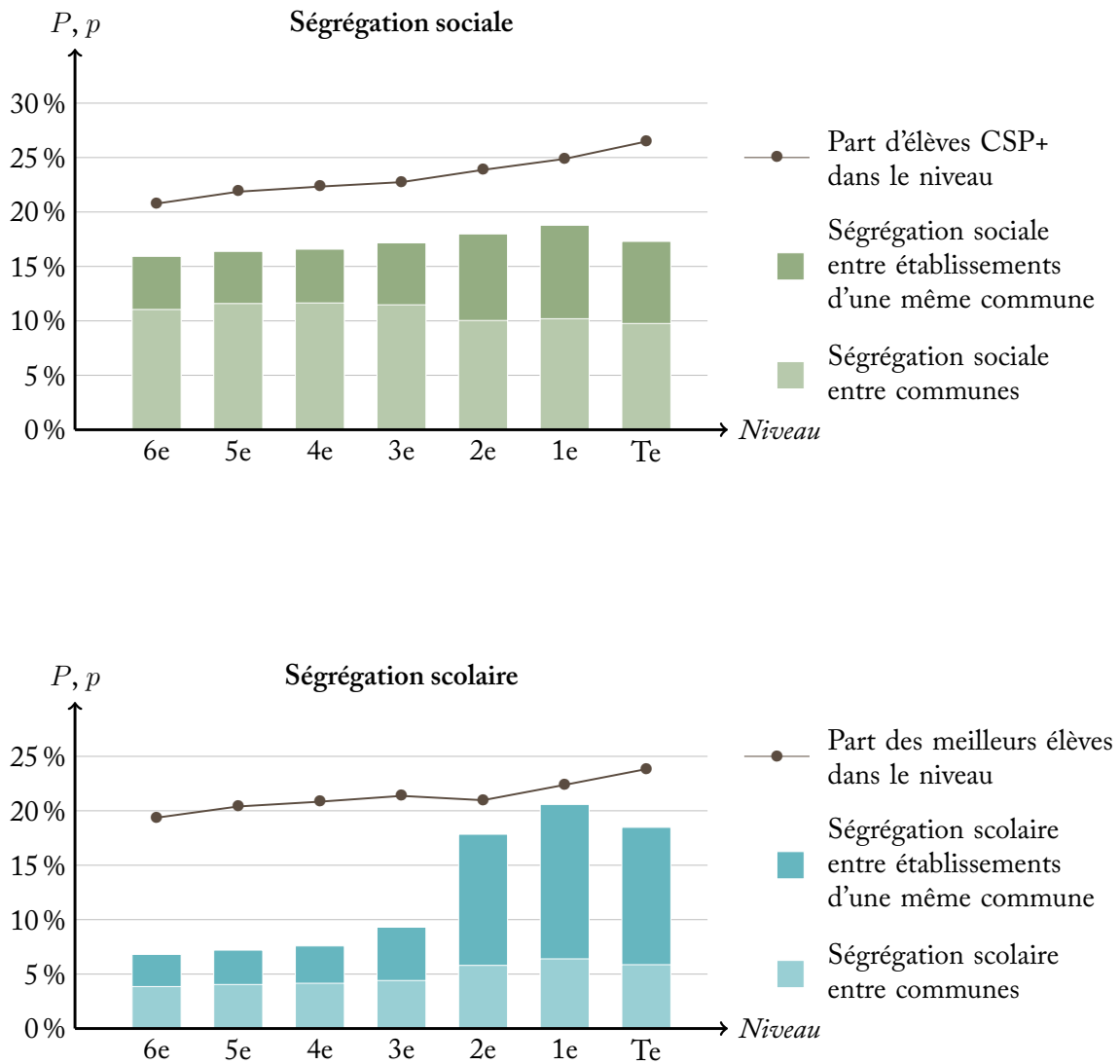


FIGURE 3.1 – Ségrégation sociale et scolaire inter-établissements au niveau national pour chaque classe, mesurée pour la cohorte 2007 par l'indice d'exposition.

Lecture : pour la cohorte d'élèves entrés en sixième en 2007, la ségrégation sociale en troisième (année 2010-2011), mesurée par l'indice d'exposition, était de 11,5 % entre communes et 17,2 % entre collèges. La part d'élèves CSP+ y était de 22,7 %.

blissements d'élèves issus de plusieurs collèges a pour effet de diminuer la ségrégation, mais la séparation des élèves entre les lycées généraux, technologiques et professionnels l'augmente. Ces deux phénomènes sont partiellement captés sur notre graphique : le premier se traduit par une baisse de la part de ségrégation expliquée par les différences entre communes, et le deuxième par une hausse de la ségrégation entre les établissements d'une même commune.

La lecture des deux composantes (inter-communes et inter-établissements, intra-communes) doit cependant être faite avec précaution. Environ 3 820 communes accueillent au moins un collège, et deux-tiers d'entre elles n'en comptent qu'un seul. Inversement, 36 % des collèges sont situés dans des communes à un seul collège, et 22 % dans des communes à deux collèges. La ségrégation sociale intra-commune, inter-établissements visible sur la **FIGURE 3.1** n'est donc générée que par un petit nombre de communes (environ 1 300).

Il faut enfin noter que la ségrégation inter-communes ne capte qu'une partie de la ségrégation résidentielle. La ségrégation intra-commune, inter-établissements, est la somme de deux composantes que nos données ne permettent pas de distinguer. D'une part, elle reflète la ségrégation entre les différents quartiers de chaque commune, chaque établissement accueillant en majorité les élèves qui habitent à proximité de celui-ci. À cette composante s'ajoutent des mouvements d'élèves qui choisissent, par dérogation ou en s'inscrivant dans le secteur privé, d'aller dans un établissement autre que leur établissement de secteur.

Nous présentons en annexe (**FIGURE 3.B.1**, en haut) les chiffres de la ségrégation sociale mesurée en prenant comme groupe de référence les élèves CSP-. Le constat est similaire : l'indice d'exposition varie entre 14 et 16 % selon le niveau.

3.3.1.2 Ségrégation scolaire

La ségrégation scolaire a une évolution beaucoup plus contrastée en fonction du niveau que la ségrégation sociale. Dans les chiffres nationaux, elle reste relativement limitée au collège, où elle varie de 7 à 9 %. Ainsi, alors qu'en l'absence de ségrégation tous les élèves compteraient 21 % des « bons élèves » parmi les élèves de même niveau dans leur collège, les bons élèves en comptent 27 % et les élèves moyens ou en difficulté n'en comptent que 19 %.

Le niveau total de ségrégation double cependant entre le collège et le lycée, pour atteindre 18 à 21 %. L'augmentation brutale de la ségrégation entre la troisième et la seconde s'explique en premier lieu par la séparation de la majorité des élèves entre des lycées généraux et technologiques et des lycées professionnels. 42 % des lycées sont des lycées généraux et/ou technologiques, 35 % sont des lycées professionnels et 20 % sont des lycées polyvalents¹⁰.

Cette rupture est beaucoup mieux repérée par le critère scolaire que par le critère social, ce qui est assez rassurant dans une certaine mesure : si très peu d'élèves parmi les meilleurs choisissent de suivre la voie professionnelle, on y trouve néanmoins une quantité non-négligeable d'élèves issus des classes aisées.

3.3.1.3 Le rôle du privé

La segmentation entre les secteurs public et privé est très souvent présentée, dans le débat public, comme la première source de ségrégation. De fait, les établissements privés accueillent, en moyenne, un public nettement plus favorisé que les établissements publics. À tous les niveaux de la scolarité secondaire, le secteur privé représente environ 21 % des élèves. Parmi eux, 32 à 37 % sont des élèves CSP+, contre seulement 18 à 24 % dans le secteur public (les chiffres varient selon le niveau, mais l'écart public-privé demeure proche de 14 points de pourcentage). Bien que plus faibles, les inégalités *scolaires* entre les deux secteurs restent importantes : la part de meilleurs élèves varie entre 26 et 29 % dans le secteur privé, alors qu'elle est comprise entre 18 et 23 % dans les établissements publics. Ainsi, on observe une différence notable entre les caractéristiques moyennes des élèves des secteurs public et privé.

Ces moyennes cachent cependant d'importantes disparités entre les établissements de chaque secteur. En effet, la séparation public-privé n'explique, au niveau national, que 8 à 13 % de la ségrégation sociale entre les établissements. Cela signifie que même si les établissements privés accueillent en moyenne un public nettement plus favorisé, les différences entre un établissement favorisé et un établissement défavorisé du même secteur (public ou privé) sont beaucoup plus importantes que les différences entre un établissement moyen public et un établissement moyen

10. Les 3 % restants correspondent essentiellement aux Établissements régionaux d'enseignement adapté (EREA, 2 %) ainsi qu'à certains collèges et autres établissements spécialisés qui contiennent des classes de seconde professionnelle.

privé. En termes de ségrégation scolaire, le rôle du privé est plus limité encore : il explique 10 à 11 % de la ségrégation entre les collèges, et seulement 1 à 2 % de la ségrégation entre les lycées.

Si ces chiffres peuvent paraître surprenants, c'est parce qu'ils comparent l'écart de composition sociale et scolaire entre les deux secteurs à toute la diversité des publics accueillis dans l'ensemble des établissements du territoire. Si les acteurs de terrain ont le sentiment que la séparation public-privé est la première source de ségrégation sociale entre établissements, c'est parce qu'elle joue un rôle beaucoup plus important à l'échelle locale. En effet, la segmentation public-privé explique 40 à 45 % de la ségrégation sociale et 33 à 40 % de la ségrégation scolaire entre les collèges *d'une même commune*. Cela signifie que les différences de composition sociale entre un collège public favorisé et un collège public défavorisé d'une même commune sont du même ordre de grandeur que les différences entre un collège public moyen et un collège privé moyen. Le rôle du privé diminue sensiblement au lycée, où la séparation entre les secteurs n'explique que 22 à 25 % de la ségrégation sociale et 13 à 14 % de la ségrégation scolaire entre établissements d'une même commune.

Il faut cependant garder à l'esprit que l'existence des deux secteurs n'est qu'un mécanisme de ségrégation parmi d'autres : la réintégration des collèges privés dans le secteur public, par exemple, n'aurait pas pour effet mécanique de réduire la ségrégation inter-établissements au sein des communes de 40 %. Une politique de réduction de la ségrégation entre établissements ne pourrait en aucun cas se limiter à une fusion des deux secteurs sans remise en question de la sectorisation et des politiques d'affectation.

3.3.2 Disparités géographiques

Les chiffres nationaux de la ségrégation sociale et scolaire peuvent paraître relativement modestes en première lecture. Ils ne sont cependant que des indicateurs synthétiques d'une situation beaucoup plus complexe.

On observe par exemple que plusieurs établissements accueillent soit un très petit nombre, soit un très grand nombre d'élèves CSP+ ou de bons élèves. Ainsi, en troisième, 10 % des élèves fréquentent des établissements contenant 5 % ou moins d'élèves CSP+ dans leur niveau ; à l'in-

verse, 5 % des élèves ont plus de 60 % d'élèves CSP+ dans leur cohorte, et même plus de 80 % de CSP+ pour les 1 % d'élèves (soit plus de 7000 élèves) dont les environnements sont les plus favorisés. Si on s'intéresse aux élèves issus des milieux les plus populaires (catégorie « Défavorisés » de la Depp, voir **TABLE 3.1**), qui représentent 37 % des élèves de troisième, 10 % des élèves en comptent 63 % ou plus dans leur établissement ; 5 % en comptent 71 % ou plus. De tels écarts sont également observés en termes de ségrégation scolaire : 10 % d'élèves comptent 6 % ou moins de « bons élèves » dans leur cohorte, et à l'inverse 5 % d'élèves en comptent plus de 43 %, et 1 % en comptent plus de 58 %.

Ces disparités sont fortement marquées géographiquement. La **FIGURE 3.2** montre les variations entre départements de l'indice de ségrégation sociale, et la **FIGURE 3.3** montre les variations de l'indice de ségrégation scolaire. La **FIGURE 3.4** donne la valeur des indices de ségrégation académique par académie. Ces trois figures portent uniquement sur le niveau troisième.

D'un département à un autre, la ségrégation sociale varie de 2 % à 27 %. Les départements où la ségrégation est plus faible sont des départements fortement ruraux (Lozère, Ariège, Lot, Aude). Dans ces départements à faible densité de population, les collèges recrutent sur un rayon pouvant dépasser les dix kilomètres : ils regroupent donc dans un même lieu des élèves d'origines différentes, ce qui favorise la mixité sociale. À l'inverse, les départements ayant la plus forte ségrégation sociale sont essentiellement des départements urbains qui comportent des grandes villes (les Hauts-de-Seine et Paris se dégagent nettement, suivis des Yvelines, du Val-de-Marne, du Nord, du Rhône et des Bouches-du-Rhône). La multiplication du nombre de collèges dans ces zones augmente au contraire la ségrégation par deux biais : d'abord, parce que les collèges reflètent plus précisément la ségrégation résidentielle, et ensuite parce qu'il s'installe une situation de concurrence qui fait émerger des collèges « souhaités » et des collèges « évités ».

Il faut rester très prudent quant à l'interprétation de ces disparités : il ne suffit en effet pas de comparer deux indices de ségrégation pour conclure qu'un territoire est plus ségrégué que l'autre, en particulier si ces territoires ont une composition sociale ou scolaire différente. Par exemple, l'académie de Paris, qui possède les indices de ségrégation les plus élevés, est aussi celle qui compte le plus d'élèves CSP+ et d'élèves parmi les meilleurs. Des comparaisons plus

raisonnables peuvent être faites entre des académies qui ont une composition sociale ou scolaire proche. Par exemple, l'académie de Lille, qui a une composition sociale proche des académies de Reims, Nancy-Metz, Amiens ou Caen, a un indice de ségrégation sociale nettement plus élevé que ces académies. Il en va de même pour l'académie de Créteil, dont l'indice d'exposition est nettement plus élevé que les académies de Strasbourg et Rennes dont les compositions sociales sont proches. En termes de ségrégation scolaire, les académies de Créteil, Lille, Nice, Versailles, Strasbourg et Paris ont des indices de ségrégation nettement supérieurs à ceux d'académies ayant des compositions scolaires comparables.

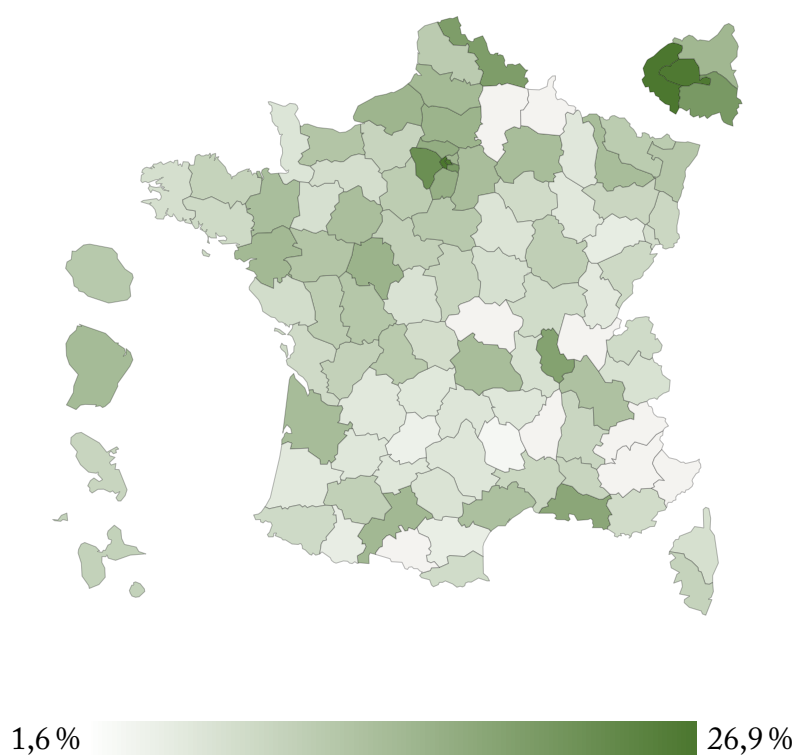


FIGURE 3.2 – Ségrégation sociale inter-établissements en classe de troisième par département, mesurée par l'indice d'exposition à la rentrée 2010.

La **FIGURE 3.5** montre par ailleurs la forte corrélation entre les niveaux de ségrégation sociale et de ségrégation scolaire de chaque département (aucun point clairement éloigné de la diagonale centrale) : les départements dans lesquels il existe une forte ségrégation sociale sont également ceux dans lesquels il existe une forte ségrégation scolaire. Cette corrélation peut notamment s'expliquer par la corrélation qui existe entre les résultats scolaires et l'origine sociale.

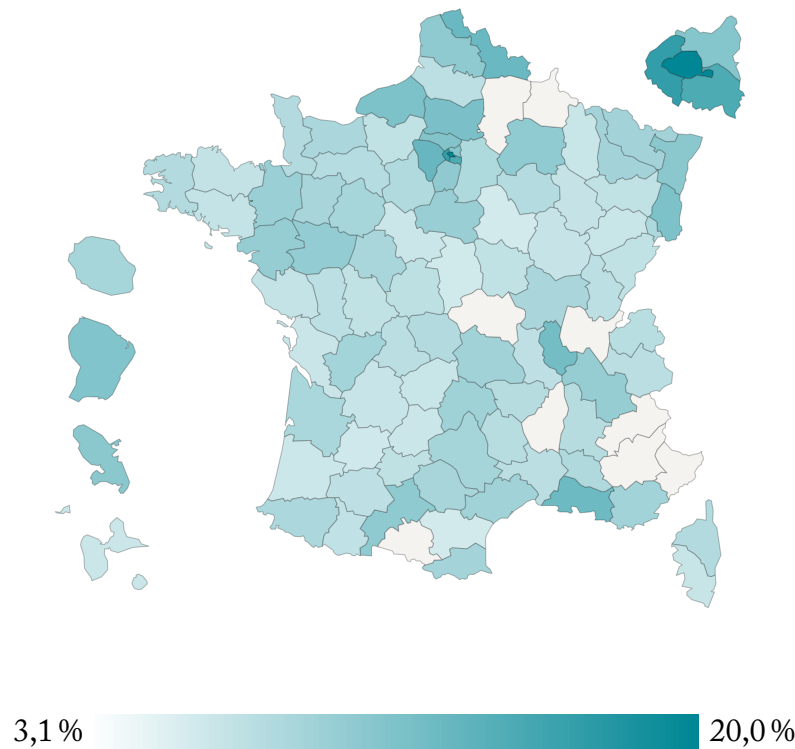


FIGURE 3.3 – Ségrégation scolaire inter-établissements en classe de troisième par département, mesurée par l'indice d'exposition à la rentrée 2010.

3.4 La ségrégation sociale et scolaire entre les classes de chaque établissement

La dimension intra-établissement est absente de toutes les études quantitatives sur la ségrégation sociale et scolaire en France. La première explication à cette carence est liée à l'accès aux données : les bases de données généralement communiquées aux chercheurs ne contiennent habituellement pas d'informations quant à la constitution des classes. Cette contrainte a pu être levée grâce à un partenariat avec la Depp, que nous remercions chaleureusement pour sa coopération et son accueil. La deuxième explication est plus technique : les outils classiques de mesure présentés dans le chapitre 3.2 présentent des faiblesses lorsque les unités qui définissent l'environnement des individus sont de petite taille : il devient alors nécessaire d'introduire de nouvelles manières de mesurer la ségrégation. Enfin, la très grande diversité des mécanismes susceptibles de créer de la ségrégation entre les classes rend difficile l'analyse des données.

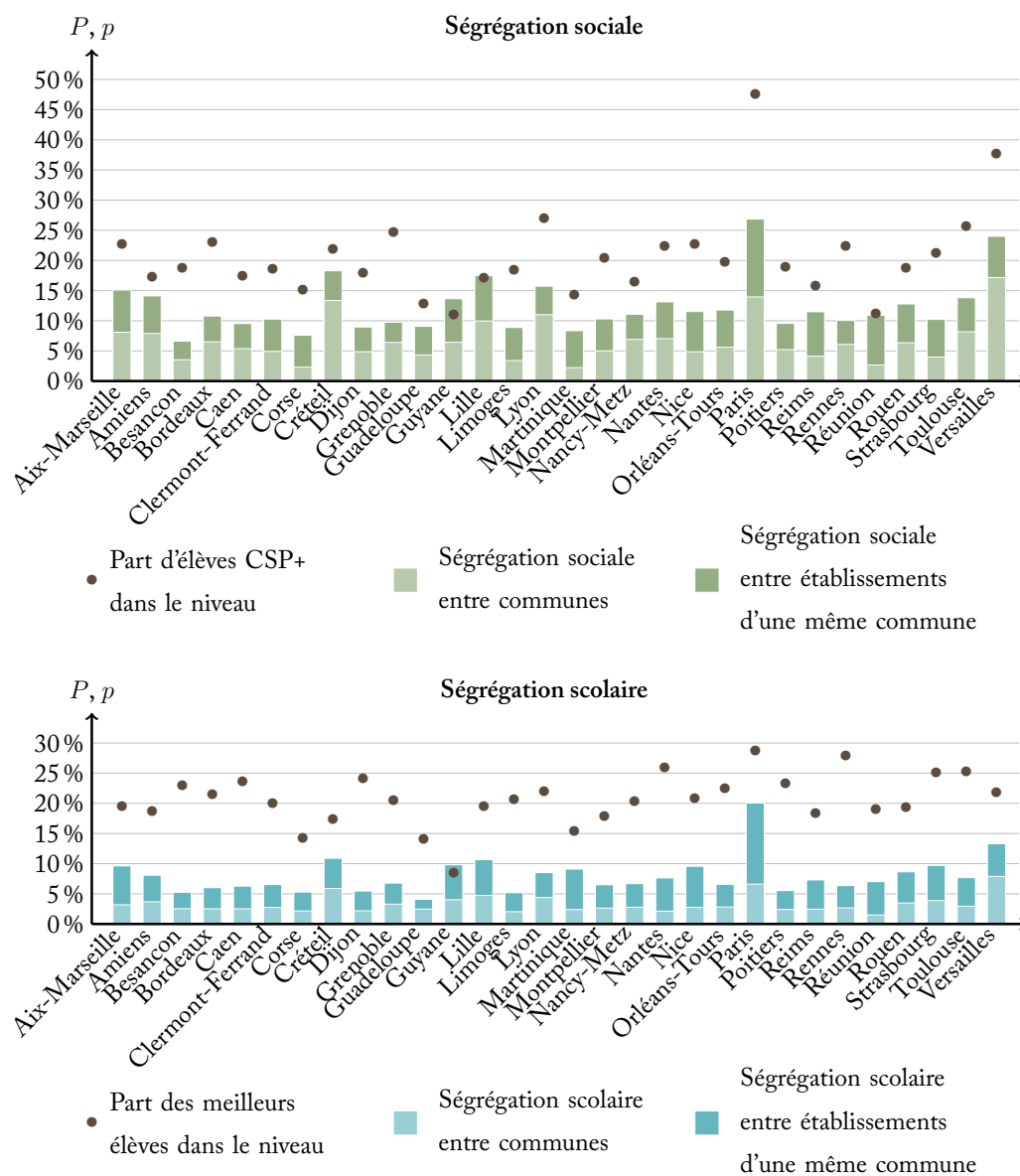


FIGURE 3.4 – Ségrégation scolaire entre établissements au niveau troisième par académie, mesurée par l'indice d'exposition à la rentrée 2010, et part expliquée par les disparités entre communes ou arrondissements municipaux.

Lecture : Dans l'académie d'Orléans-Tours, au niveau troisième, la ségrégation sociale entre établissements est de 12 %, dont 6 points s'expliquent par les disparités entre communes. Dans l'académie de Versailles, la ségrégation scolaire entre établissements vaut 13 %, dont 8 points s'expliquent par les disparités entre communes.

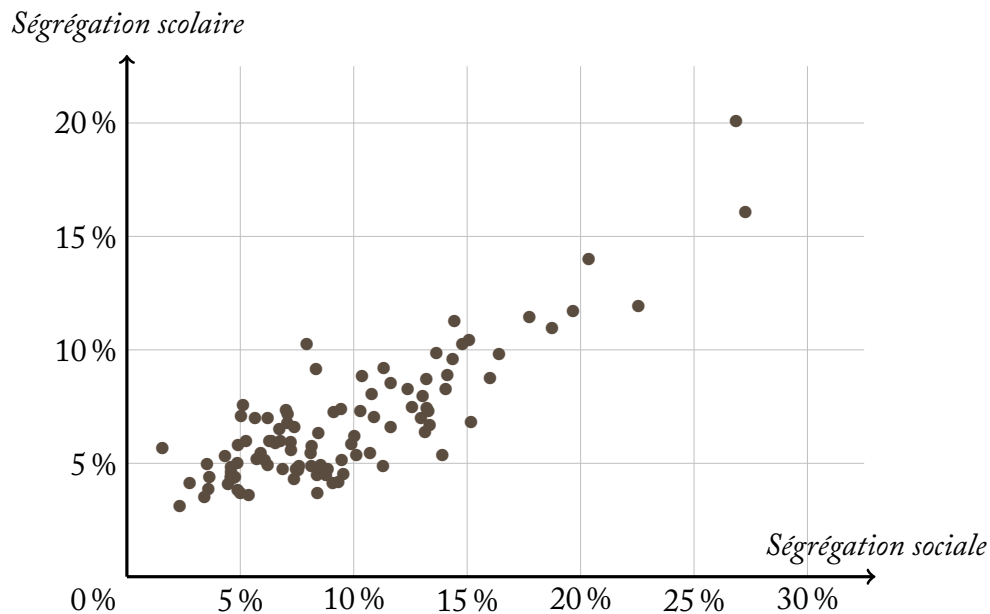


FIGURE 3.5 – Corrélation entre la ségrégation sociale et la ségrégation scolaire dans chaque département au niveau troisième, pour l'année 2010-2011.

Chaque point représente un département : son abscisse (axe horizontal) représente sa ségrégation sociale et son ordonnée (axe vertical) représente sa ségrégation scolaire, les deux étant mesurées par l'indice d'exposition. On observe que les points sont globalement très proches de la diagonale, ce qui indique que les deux formes de ségrégation sont très corrélées.

La prise en compte de la ségrégation intra-établissement est pourtant primordiale. Pour fixer les idées, considérons deux collèges A et B, contenant chacun deux classes de mêmes tailles. Supposons que la situation initiale est très ségrégée, et que le collège A contient 80 % de CSP+ alors que le collège B n'en contient que 20 %. Dans chacun des collèges, on peut imaginer que l'homogénéité des élèves n'incite pas à mettre en place une forte ségrégation entre les classes, et donc que chaque classe contient la même part de CSP+ qu'il y en a dans le collège. Il y a donc, dans ce scénario, deux classes à 80 % de CSP+ dans le collège A, et deux classes à 20 % de CSP+ dans le collège B. Imaginons alors qu'une politique très volontariste de mixité force le mélange de ces deux établissements, de sorte qu'ils contiennent chacun 50 % de CSP+. Chaque collège accueille un public beaucoup plus hétérogène, ce qui peut créer des tensions et une demande de différenciation des classes. Dans un scénario pessimiste, on pourrait ainsi aboutir à deux collèges contenant chacun une classe à 80 % de CSP+ et une classe à 20 % de CSP+.

Si on se limite à une analyse au niveau inter-établissements, la politique de mixité donnée dans cet exemple fictif a été un franc succès, car elle a réduit la ségrégation sociale entre établissements à zéro. Pourtant, la ségrégation totale (inter- et intra-établissement), n'a pas varié : on retrouve bien deux classes à 80 % et deux classes à 20 % de CSP+. De plus, chaque élève devient très conscient qu'il appartient à la « bonne » classe ou à la « mauvaise » classe de son collège, ce qui est susceptible de créer plus de tensions encore que la situation initiale. Une politique de mixité sociale ne doit donc pas être pilotée sans indicateurs de mixité sociale au niveau intra-établissement.

Dans cette étude, nous commençons par mesurer le niveau de ségrégation qui est observé lorsque l'on s'intéresse aux environnements « classe » au lieu des environnements « établissement », afin de les comparer aux valeurs obtenues dans le chapitre 3.3. Les résultats ainsi obtenus représentent la ségrégation *totale*, qui est la somme de la ségrégation *inter-établissements* étudiée dans le chapitre 3.3 et de la ségrégation *intra-établissement*. Nous montrons ensuite les limites des indicateurs classiques pour étudier le phénomène de la ségrégation intra-établissement en raison de la petite taille des unités de séparation (les classes) et nous proposons des méthodes d'analyse supplémentaires. Nous nous intéressons enfin plus précisément au rôle des langues vivantes et des options dans la composition des classes.

3.4.1 Analyse macroscopique

La **FIGURE 3.6** donne le niveau de ségrégation totale (ségrégations inter- et intra-établissement cumulées) à chaque niveau, dans les dimensions sociale et scolaire. Les niveaux de ségrégation inter-établissements y sont rappelés afin de distinguer les composantes inter- et intra-établissement.

La ségrégation sociale augmente de 4 à 6 points en fonction du niveau lorsque la ségrégation intra-établissement est prise en compte, pour atteindre 20 à 22 points au collège et en seconde, 25 points en première et 23 points en terminale. Ainsi, un collégien CSP+ compte en moyenne 5 élèves eux-mêmes CSP+ dans sa classe de plus qu'un collégien non-CSP+ (une classe de collège comporte 25 élèves en moyenne). La proportion d'élèves CSP+ au collège étant d'environ 5 par classes, cela signifie qu'en moyenne, *les collégiens CSP+ comptent deux fois plus de camarades de*

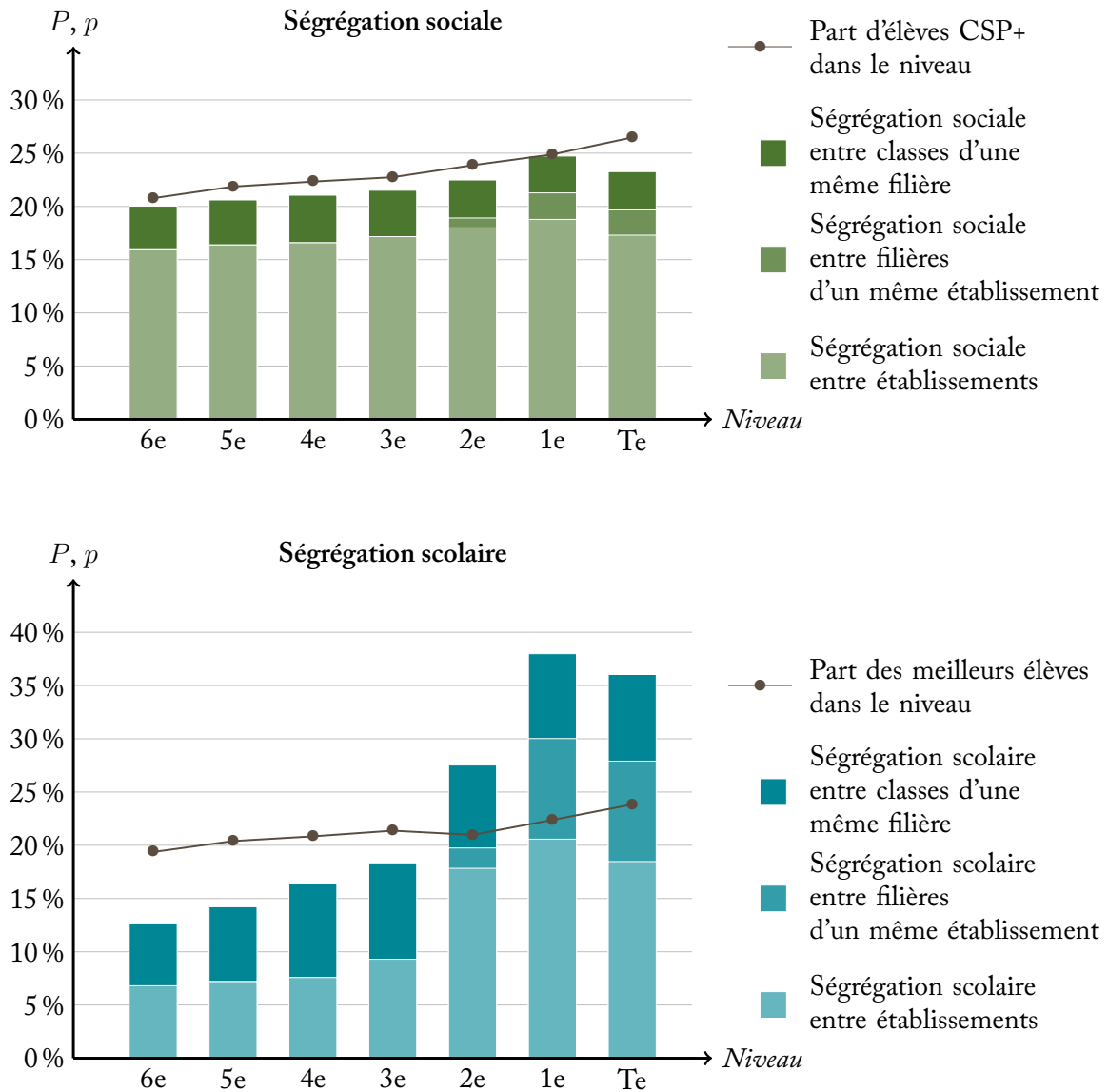


FIGURE 3.6 – Ségrégation sociale et scolaire inter- et intra-établissement au niveau national pour chaque classe, mesurée pour la cohorte 2007 par l'indice d'exposition.

Lecture : pour la cohorte d'élèves entrés en sixième en 2007, la ségrégation scolaire en première (année 2012-2013), mesurée par l'indice d'exposition, était de 20,6 % entre établissements, et 30,0 % entre filières et 38,0 % entre classes. La part des meilleurs élèves y était de 22,4 %.

classe CSP+ que les autres collégiens (non-CSP+). Il est important de noter que ce résultat n'est pas le seul fait de la composition des classes : au contraire, il s'explique à 80 % par la seule ségrégation inter-établissement. L'effet additionnel de la composition des classes sur la ségrégation sociale est donc relativement modeste si on considère ces chiffres moyens.

La ségrégation scolaire, elle, est beaucoup plus fortement portée par la dimension intra-établissement : au collège, les dimensions inter-établissements et intra-établissement ont une contribution quasi-identique à la ségrégation totale, c'est-à-dire que *les classes de niveau contribuent autant à la ségrégation scolaire que les disparités résidentielles*. Cette ségrégation augmente progressivement au cours de la scolarité au collège, pour atteindre 18 % en troisième ¹¹. La conclusion sur la ségrégation scolaire totale au collège est ainsi identique que pour la ségrégation sociale : *un très bon élève sera entouré, dans sa classe, de deux fois plus d'élèves d'un niveau équivalent au sien que les autres élèves*. La structure de cette ségrégation est cependant différente, puisqu'elle résulte pour 50 % de la ségrégation inter-établissements et pour 50 % de la ségrégation intra-établissement.

Au lycée, le niveau de ségrégation scolaire devient très élevé et atteint 38 % en première. Il est à noter que cette très forte augmentation s'explique en grande partie par les différences de niveau des élèves en fonction de leurs voies et séries du baccalauréat ou du CAP. En effet, la dimension intra-établissement inclut ainsi la séparation des élèves entre les voies générale et technologique à partir de la première dans les lycées « GT » (généraux et technologiques), ainsi que la séparation avec les élèves de la voie professionnelle dans les lycées polyvalents. Dans tous les cas, elle inclut également le sectionnement des élèves en fonction de leur série du baccalauréat.

L'analyse des disparités géographiques montre que la ségrégation intra-établissement est un phénomène beaucoup moins variable d'un territoire à un autre. Cela se traduit par un plus faible contraste dans les cartes de la **FIGURE 3.7** (ségrégation sociale totale) et de la **FIGURE 3.8** (ségrégation scolaire totale).

Précisons ce constat numériquement en analysant la variabilité de la ségrégation intra-établissement entre les académies. On obtient la ségrégation intra-établissement de chaque

11. Rappelons que la valeur de l'indice de ségrégation scolaire est également mesurée plus précisément en troisième que dans les autres classes.

académie en soustrayant aux chiffres de la ségrégation totale ceux de la ségrégation inter-établissements : cela correspond à la zone la plus foncée de chaque graphe sur la **FIGURE 3.9**. On constate tout d'abord que les académies d'outre-mer se caractérisent par une ségrégation scolaire intra-établissement beaucoup plus forte que les académies métropolitaines.

En se restreignant aux académies métropolitaines, la ségrégation sociale inter-établissements vaut en moyenne 14,1 % avec un écart-type entre les académies de 4,8 points, et la ségrégation scolaire inter-établissements vaut 8,7 % avec un écart-type de 3,2 points : il existe donc des disparités importantes entre les académies sur le plan de la ségrégation inter-établissement. Dans la dimension intra-établissement, la valeur moyenne de la ségrégation sociale d'une académie est plus faible (4,5 %), et son écart-type qui n'est que de 0,6 points. La ségrégation scolaire intra-établissement reste élevée (8,9 % en moyenne interacadémique), mais son écart-type n'est que d'1,1 point. Ainsi, *l'ensemble des académies métropolitaines se caractérisent par une ségrégation intra-établissement modeste dans la dimension sociale et aussi importante que la ségrégation inter-établissements dans la dimension scolaire*. Dans les académies ultramarines, la ségrégation scolaire intra-établissement dépasse même la ségrégation scolaire inter-établissements.

3.4.2 Mieux mesurer la ségrégation intra-établissement

Les valeurs de ségrégation intra-établissement présentées dans la section précédente doivent être interprétées avec précaution. Lorsque les unités qui définissent l'environnement des élèves sont de petite taille (ici, les classes comportent en général moins de 30 élèves), des écarts d'un ou deux élèves ont un effet sur les indicateurs de plusieurs points de pourcentage. Dès lors, une ségrégation intra-établissement de 5 % – l'ordre de grandeur de la ségrégation sociale intra-établissement – ne doit pas être interprétée comme le résultat d'une politique d'établissement qui recherche activement à ségréger.

Un chef d'établissement qui ne chercherait pas spécialement à faire des classes ségréguées, par exemple en les constituant totalement aléatoirement, pourrait aboutir à une ségrégation intra-établissement du même ordre de grandeur. Même en essayant de faire moins ségrégué que l'aléatoire, il serait difficile (ou impossible) d'obtenir des classes parfaitement équilibrées, surtout si on cherche à équilibrer selon plusieurs dimensions (sociale, scolaire, mixité de genre, d'origine

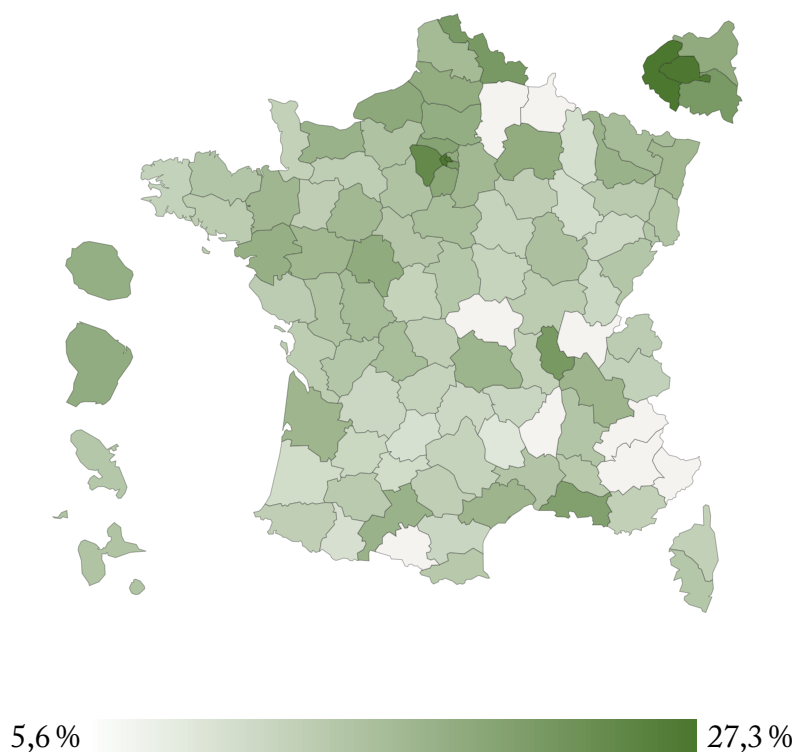


FIGURE 3.7 – Ségrégation sociale totale (inter- et intra-établissements) en classe de troisième par département, mesurée par l'indice d'exposition à la rentrée 2010.

ethnique, etc.) : une ségrégation résiduelle non négligeable apparaîtra tout simplement à cause des arrondis qu'il aura fallu effectuer, un élève ne pouvant bien entendu pas être divisé entre plusieurs classes !

Pour confirmer cette intuition, nous avons simulé des constitutions de classe aléatoires, en tirant au sort, à cent reprises, la classe de chaque élève dans son collège. Nous nous limitons au collège à ce stade car il s'agit des niveaux où la marge de manœuvre est plus grande : au lycée, les voies et les séries limitent les possibilités de « mélanger » les classes des élèves.

Pour chaque collège et chaque niveau, nous obtenons ainsi autant d'indices de ségrégation sociale et scolaire intra-établissement que de simulations. La ségrégation sociale intra-établissement, d'une valeur réelle en troisième de 5,0 % en moyenne, atteint 3,1 % en moyenne sur cent simulations. La ségrégation scolaire, qui vaut 9,6 % en réalité, n'atteint que 3,2 % dans

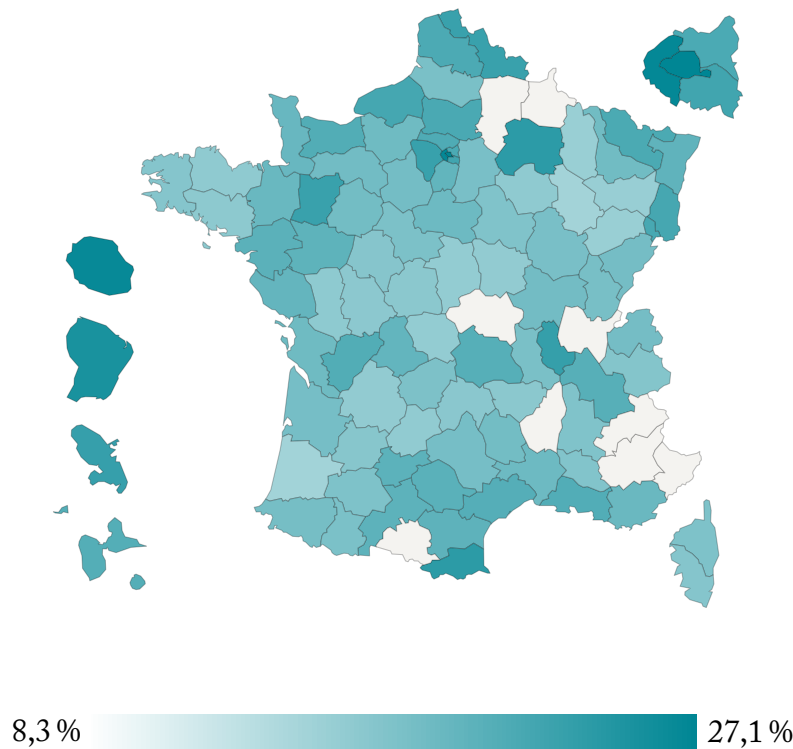


FIGURE 3.8 – Ségrégation scolaire totale (inter- et intra-établissements) en classe de troisième par département, mesurée par l'indice d'exposition à la rentrée 2010.

les simulations. Ainsi, si la ségrégation effectivement mesurée dépasse celle qui découlerait d'une répartition aléatoire des élèves, il reste difficile d'interpréter le niveau réel comme faible ou élevé.

Pour ce faire, nous reproduisons ici à l'échelle nationale le travail que nous avons effectué sur la région Île-de-France (Ly et al., 2014), qui consiste à comparer le degré de ségrégation de chaque collège à la distribution des indices de ségrégations obtenus dans les simulations¹². Pour un établissement donné, si moins de 10 % des simulations aboutissent à un niveau de ségrégation au moins égal au niveau réel, on conclut que cet établissement exerce une ségrégation « active », qui va au-delà du hasard.

Selon ce critère, en classe de troisième par exemple, on observe une ségrégation sociale allant

12. Carrington and Troske (1997) utilisent une approche similaire pour définir des indices de ségrégation « corrigés », que nous n'utilisons pas dans cette étude. Selon Rathelot (2012), les indices ainsi définis sont sous-estimés.

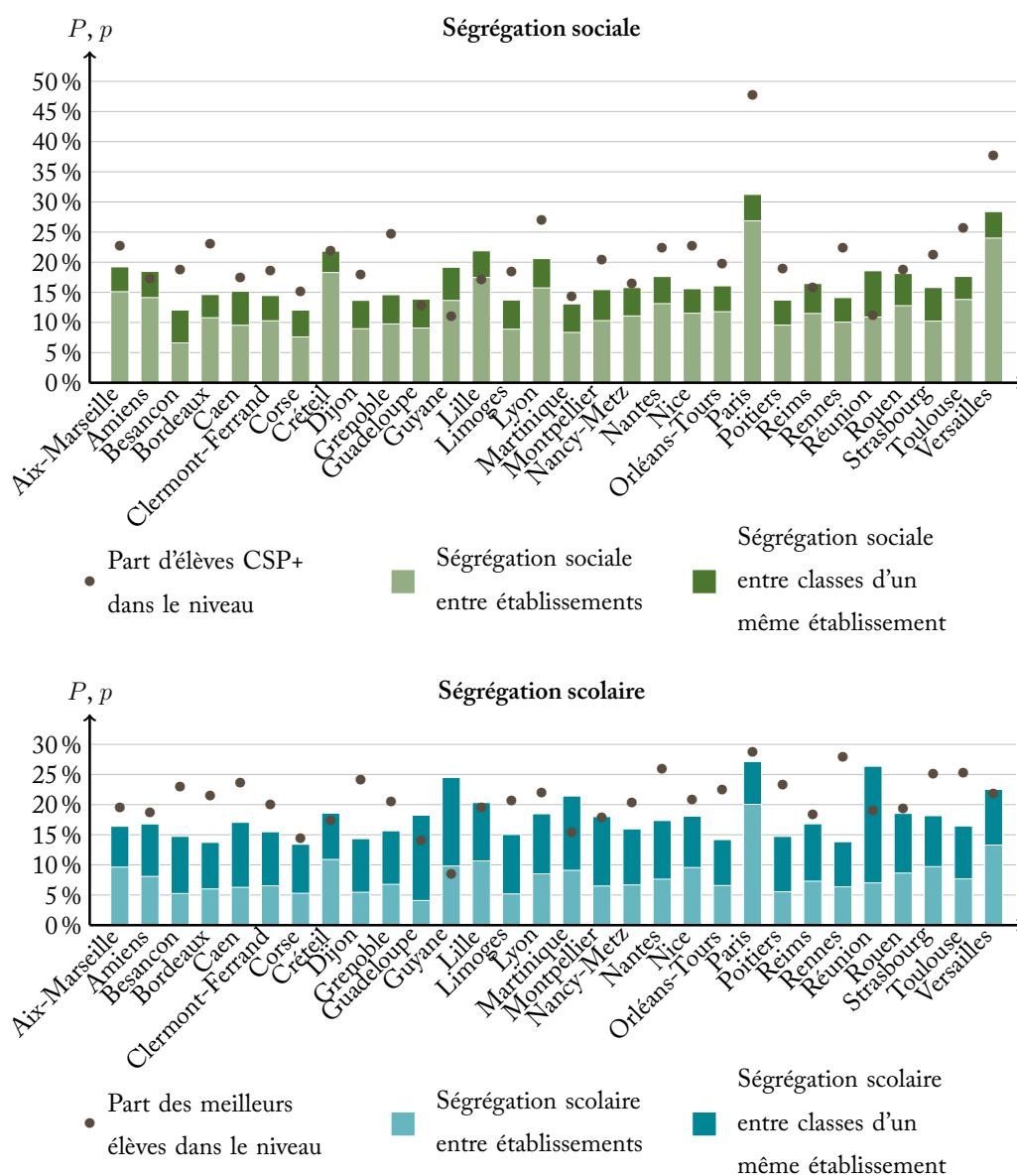


FIGURE 3.9 – Ségrégation scolaire entre classes au niveau troisième par académie, mesurée par l'indice d'exposition à la rentrée 2010, et part expliquée par les disparités entre établissements. Lecture : Dans l'académie d'Orléans-Tours, au niveau troisième, la ségrégation sociale entre classes est de 16 %, dont 12 points s'expliquent par les disparités entre établissements. Dans l'académie de Versailles, la ségrégation scolaire entre communes vaut 23 %, dont 13 points s'expliquent par les disparités entre établissements.

au-delà du hasard dans 25 % des collèges, et une ségrégation scolaire plus élevée que le hasard dans 45 % des collèges. Il faut bien noter qu'une ségrégation « active » selon notre critère n'est pas nécessairement « élevée » : ce critère signifie simplement qu'on ne peut pas expliquer les disparités entre classes uniquement par le hasard de la composition des classes. Il est alors important d'analyser également le chiffre brut de la ségrégation intra-établissement et de discuter son ampleur. Ainsi, en troisième, les 25 % de collèges qui exercent une ségrégation sociale qui dépasse l'aléatoire ont une ségrégation intra-établissement moyenne de 9 %, un chiffre équivalent à la moitié de la ségrégation inter-établissements à ce niveau. Toujours en troisième, les 45 % de collèges qui exercent une ségrégation scolaire « active » ont une ségrégation intra-établissement moyenne de 14 %, soit une fois et demie la ségrégation inter-établissements à ce niveau.

Une technique supplémentaire consiste à regarder, au lieu de l'indice d'exposition, la composition de la « meilleure classe » qui résulte de chaque allocation aléatoire des élèves. On peut alors comparer cette composition à celle de la « vraie » meilleure classe de chaque collège. Cette méthode permettrait de détecter certains établissements qui veilleraient à bien équilibrer la majorité des classes, à l'exception d'une classe très favorisée. Pour cela, nous calculons pour chaque simulation la part d'élèves « privilégiés » (CSP+ ou meilleurs élèves) dans la classe qui en contient le plus. Nous mesurons ensuite la part d'élèves privilégiés dans la classe réelle qui en contient le plus, et nous comparons cette valeur réelle aux valeurs aléatoires. Là encore, si la valeur réelle n'est dépassée que par 10 % ou moins des valeurs obtenues par les simulations, on considère que l'établissement a activement ségrégué les élèves en créant une classe nettement plus favorisée que les autres.

Ce deuxième critère fournit des résultats semblables au premier, que nous ne détaillons pas ici. Plus de 80 % des établissements « activement ségrégués » selon l'un des critères le sont également selon l'autre critère.

Les deux méthodes proposées ici sont des premières pistes pour analyser la ségrégation intra-établissement ; nous souhaitons en proposer d'autres et nous invitons les chercheurs, décideurs et acteurs de terrain à participer à cette réflexion.

3.4.3 Le rôle des langues et des options

Les langues vivantes et les options sont souvent vues, dans le débat public, comme des outils permettant de ségréger les classes. Elles seraient utilisées par les familles pour accéder aux « meilleures classes », et par les chefs d'établissement pour éviter la fuite des classes moyennes et supérieures vers les établissements concurrents, en particulier les établissements privés dans les quartiers populaires (Baluteau, 2013).

Nous nous intéressons dans cette étude à trois options réputées élitistes que sont le parcours bilangues en sixième et cinquième, l'option latin à partir de la cinquième et la section européenne à partir de la quatrième. Nous nous limitons, pour chaque option, au niveau à partir duquel elle apparaît. Dans cette partie, qui ne porte que sur le collège, nous étudierons la cohorte d'élèves entrés en sixième en 2010 et non en 2007 comme dans le reste de l'étude¹³. Il est en effet important d'utiliser des données aussi récentes que possible, car la place des options a fortement évolué au cours des dernières années, en raison de l'assouplissement de la carte scolaire en 2007, mais aussi d'évolutions tendanciennes de disponibilité et de popularité des différentes options. Par exemple, depuis le début des années 2000, les données à notre disposition montrent que les parcours bilangues et les sections européennes ont progressé fortement mais régulièrement – sans accélération après l'assouplissement de la carte scolaire. Alors que ces deux options ne concernaient que 2 % des élèves de sixième et quatrième à la rentrée 2001, les parcours bilangues étaient choisis par 16 % des élèves de sixième et les sections européennes par 11 % des élèves de quatrième à la rentrée 2014. Dans le même temps, l'option latin en cinquième a légèrement décliné, de 22 % à 20 %.

3.4.3.1 Le parcours bilangues

À la rentrée 2010, 13 % des élèves de sixième étudiaient l'anglais et une deuxième langue vivante, l'allemand dans 72 % des cas¹⁴. Parmi l'ensemble des germanistes, 82 % appartenaient

13. Dans le reste de l'étude, nous avons choisi la cohorte 2007 car il s'agit de la dernière cohorte étant sortie du lycée pour laquelle nous disposons des données nécessaires à ce jour.

14. Nous excluons du cadre de cette étude les langues régionales, qui ne sont choisies que par 1 % des élèves. L'apprentissage du corse est quasi-généralisé en Corse, ce qui n'en fait pas un outil de ségrégation. Les deux autres langues régionales les plus enseignées sont le breton et l'occitan ; de manière générale, nous n'observons pas un fort marquage social ou scolaire lié à l'apprentissage d'une langue régionale.

à une classe bilangues.

Cette option est marquée socialement (33 % des élèves concernés sont CSP+, alors que ceux-ci représentent 22 % des élèves de sixième) et surtout scolairement : 38 % des élèves des classes bilangues font partie de la catégorie des meilleurs élèves, qui ne représente que 20 % des élèves de sixième.

Le parcours bilangues est disponible dans 53 % des collèges, soit 3 707 établissements. Cette proportion augmente régulièrement depuis le début des années 2000. Elle est fortement variable d'une académie à l'autre, allant de 23 % dans l'académie de Lille¹⁵ jusqu'à 99 % dans l'académie de Strasbourg (où l'apprentissage de l'allemand est prépondérant).

Cette valeur intermédiaire fait du parcours bilangues un outil de différenciation qualitatif *entre établissements*. Cette différenciation ne se traduit cependant pas par des écarts importants en termes de composition sociale et scolaire entre les établissements proposant des parcours bilangue et les autres : en moyenne, la part d'élèves CSP+ ou la part d'élèves parmi les meilleurs ne diffère pas de plus de 2 % entre ces deux catégories d'établissements. Ce faible écart n'exclue pas que les options jouent un rôle important dans la ségrégation entre établissements (voir la conclusion de cette section).

Nous nous concentrons dans cette étude au parcours bilangues comme outil de ségrégation entre les classes de chaque établissement. Au sein des établissements proposant un parcours bilangues, seules 42 % des classes contiennent des élèves ayant suivi un parcours bilangues. Les parcours bilangues génèrent donc une dichotomie à la fois entre les établissements et entre les classes des établissements, dans la mesure où ils existent dans environ la moitié des établissements, puis dans environ la moitié des classes de ces établissements. Parmi les établissements proposant un parcours bilangue, seuls 13 % ont réparti les élèves suivant ces parcours entre l'ensemble des classes de sixième. À l'inverse, 45 % les ont regroupés dans le minimum de classes possible¹⁶.

15. Nous excluons la Corse où la très grande majorité des élèves apprend le corse, et Mayotte où aucun élève n'est inscrit en parcours bilangues d'après les données dont nous disposons.

16. Dans deux-tiers des cas, il s'agit d'une seule classe.

Une nuance peut être apportée à ce constat : le terme couramment utilisé de « classe bilangues » peut prêter à confusion. S'il existe souvent, au sein des établissements, des classes proposant le parcours bilangues et d'autres ne la proposant pas, il est rare que tous les élèves de ces classes suivent un tel parcours. Seules 16 % de ces classes comptent plus de 80 % d'élèves suivant un parcours bilangues : les classes dites « bilangues » sont plus souvent des classes « à section bilangues », qui contiennent en moyenne 52 % d'élèves suivant un tel parcours.

Du fait du marquage social et scolaire des parcours bilangues, les classes à section bilangues sont sensiblement plus favorisées socialement (25 % de CSP+) et scolairement (26 % de la catégorie des meilleurs élèves) que les classes sans section bilangues des mêmes établissements (19 % et 16 %, respectivement). Ces écarts représentent environ un écart-type de la distribution intra-établissement des environnements de classe.

3.4.3.2 L'option latin

L'option latin est proposée dans la très grande majorité des collèges en classe de cinquième (6 509 établissements à la rentrée 2011, soit 93 %) ; elle n'est cependant choisie que par 21 % des élèves. Cette option est très fortement marquée socialement et scolairement : 37 % des latinistes sont des CSP+ (qui représentent 23 % des élèves de cinquième) et 44 % font partie des meilleurs élèves (21 % des élèves de cinquième).

Cette disponibilité quasi-générale de l'option latin n'en fait pas un fort outil de différenciation et de ségrégation *entre établissements*. Au sein des établissements qui proposent cette option, 58 % des classes contiennent des élèves latinistes. Il s'agit donc d'un facteur de différenciation des classes.

Cependant, contrairement à l'option bilangues, les élèves latinistes sont, le plus souvent, partiellement répartis entre les classes : on ne compte que 18 % des établissements qui les regroupent dans le moins de classes possible. Dans 23 % des cas, ils sont même répartis entre toutes les classes de cinquième. Il est par ailleurs très rare de trouver des classes de cinquième comportant une majorité d'élèves latinistes : seules 3 % des classes latinistes le sont à plus de 80 %.

Ces différences suffisent pourtant à générer des différences sociales et scolaires nettes entre les classes proposant l'option latin et celles ne la proposant pas. Au sein des établissements qui proposent l'option latin, les classes latinistes comptent 26 % d'élèves CSP+ et 26 % d'élèves parmi les meilleurs, alors que les autres en comptent respectivement 17 % et 15 %, soit une segmentation comparable à celle observée entre les classes à section bilingue et les autres en sixième.

3.4.3.3 Les sections européennes

À la rentrée 2012, 11 % des élèves de quatrième étaient inscrits dans une section européenne. Cette option consiste à suivre certains cours du tronc commun dans une langue étrangère, le plus souvent l'anglais¹⁷. Il s'agit là aussi d'une option très marquée socialement et scolairement : 40 % des élèves de section européenne sont CSP+ (contre 23 % parmi l'ensemble des élèves de quatrième) et 49 % font partie des meilleurs élèves (25 % de l'ensemble des quatrièmes).

Cette option n'est proposée que dans 44 % des collèges (3 100 établissements), ce qui en fait un facteur de différenciation qualitative entre établissements. À quelques exceptions près, cette proportion varie assez peu entre les académies. Il y a donc une offre scolaire relativement homogène sur le territoire pour cette option, et donc un potentiel quasi-généralisé de concurrence locale entre établissements. Comme pour les classes bilingues, cette différenciation n'est cependant pas associée à des différences très importantes de composition scolaire entre les établissements à section européenne et les autres.

Au sein des établissements proposant cette option, seule une classe sur deux contient une section européenne, ce qui en fait également un outil de ségrégation intra-établissement. 26 % des collèges regroupent les élèves de section européenne dans le minimum de classes et 15 % des collèges les répartissent entre toutes les classes de quatrième. Pour autant, seules 3 % des classes à section européennes contiennent plus de 80 % d'élèves suivant cette option.

Ces regroupements conduisent à une hiérarchie sociale et scolaire dans les établissements proposant l'option européenne, entre les classes à section européenne d'une part (26 % de CSP+,

17. Parmi les élèves inscrit en section européenne, 65 % la suivent en anglais, 15 % en espagnol, 14 % en allemand, 4 % en italien, les 5 % restants sont répartis entre des langues plus rares.

26 % d'élèves parmi les meilleurs) et les autres classes (respectivement 17 % et 15 %). On retrouve des différences d'un ordre de grandeur comparable – et même légèrement supérieur – à celles observées en sixième et cinquième avec les parcours bilangues et l'option latin.

3.4.3.4 Conclusion

L'analyse statistique présentée ici montre que les langues et les cours optionnels sont un outil très important de différenciation entre établissements (à l'exception du latin qui est presque systématiquement proposé) et surtout entre classes, et confirment les études qualitatives sur le sujet (Baluteau, 2013). Les faibles écarts de composition sociale et scolaire entre les établissements proposant les parcours bilangues ou les sections européennes ne contredisent pas le caractère différenciant de ces options. Comme le remarquait Baluteau (2013), on trouve souvent une abondance de cours optionnels dans les établissements les plus privilégiés, mais également dans des établissements moins favorisés et en situation de forte concurrence, qui cherchent à retenir un public plus aisé.

Au sein des établissements proposant une de ces options, on distingue trois cas de figures. Une minorité d'établissements regroupe les élèves choisissant ces options réputées élitistes dans une seule ou deux classes, qui sont bien souvent les classes les plus favorisées socialement et scolairement. À l'inverse, certains établissements, également minoritaires, font l'effort de répartir les élèves entre toutes les classes, ce qui contribue à l'équilibre entre celles-ci. Entre ces deux scénarios, la majorité des établissements effectue un regroupement partiel, créant des sections au sein de chaque classe correspondant aux différentes options, mais dans quelques classes seulement. Quelles que soient les motivations derrière ces pratiques (contraintes d'emploi du temps ou désir réel de construire des classes spécialisées), elles conduisent à une hiérarchie sociale et scolaire où les élèves sont en mesure de reconnaître les classes privilégiées des autres classes.

Ce constat pose la question de l'opportunité du maintien d'un large catalogue d'options au collège, du moins dans leur forme actuelle. Si l'on peut souhaiter l'existence d'une certaine liberté de choix dans les cours suivis par les élèves, le caractère optionnel de cette offre, sa disponibilité variable entre et au sein des territoires, et le fait qu'elle structure la composition des établissements et des classes favorisent les pratiques ségrégatives. À l'inverse, il n'est pas certain que la

suppression pure et simple d'un certain nombre d'options soit une mesure suffisante pour limiter la ségrégation sociale et scolaire dans les établissements, celle-ci pouvant réapparaître à terme sous d'autres formes. Surtout, cette suppression pourrait accentuer la ségrégation territoriale, si elle augmente la propension des familles les plus aisées (ou les moins défavorisés) d'un territoire à migrer vers des zones toujours plus favorisées (selon le mécanisme de séparatisme social décrit par Maurin, 2004).

3.5 Évolution de la ségrégation

Les chiffres présentés jusqu'ici concernent uniquement la cohorte d'élèves entrés en sixième en 2007. Nous nous intéressons dans cette partie à l'évolution de ces chiffres sur une période de neuf ans¹⁸. Nous avons choisi de nous concentrer sur les classes de sixième (FIGURE 3.10), troisième (FIGURE 3.11) et seconde (FIGURE 3.12). Pour chacune de ces classes, nous indiquons la valeur de la ségrégation sociale et scolaire entre établissements et entre classes d'un même établissement, pour les cohortes d'élèves ayant passé le DNB entre 2006 et 2014.

La comparaison d'indices de ségrégation entre des territoires ou des périodes différentes est délicate. Une valeur de l'indice de ségrégation n'a pas la même signification en termes d'interactions potentielles des élèves en fonction du contexte et notamment de la part du groupe de référence dans la population (voir annexe 3.A.2). Pour cette raison, nous indiquons systématiquement la proportion du groupe de référence dans l'échantillon dont nous mesurons la ségrégation sociale ou scolaire.

Dans cette section, nous avons cherché à déterminer le sens de l'évolution de la ségrégation sociale et scolaire sur une période de 9 années. Il s'agit de déterminer si les variations de l'indice de ségrégation observées d'une année sur l'autre reflètent une réelle tendance à la hausse ou à la baisse, ou simplement des écarts statistiques autour d'une tendance stable. Pour cela, nous avons calculé l'évolution moyenne de chaque indice de ségrégation, et nous l'avons comparée avec l'« erreur-type » de ces indices, c'est-à-dire la précision avec laquelle ces indices sont mesurés¹⁹.

18. Nous disposons des résultats au DNB entre 2006 et 2014, nous mesurons donc l'évolution de la ségrégation sur les cohortes d'élèves ayant passé le DNB l'une de ces années.

19. Pour mesurer l'erreur-type des indices de ségrégation, nous avons effectué 25 ré-échantillonnages avec remise

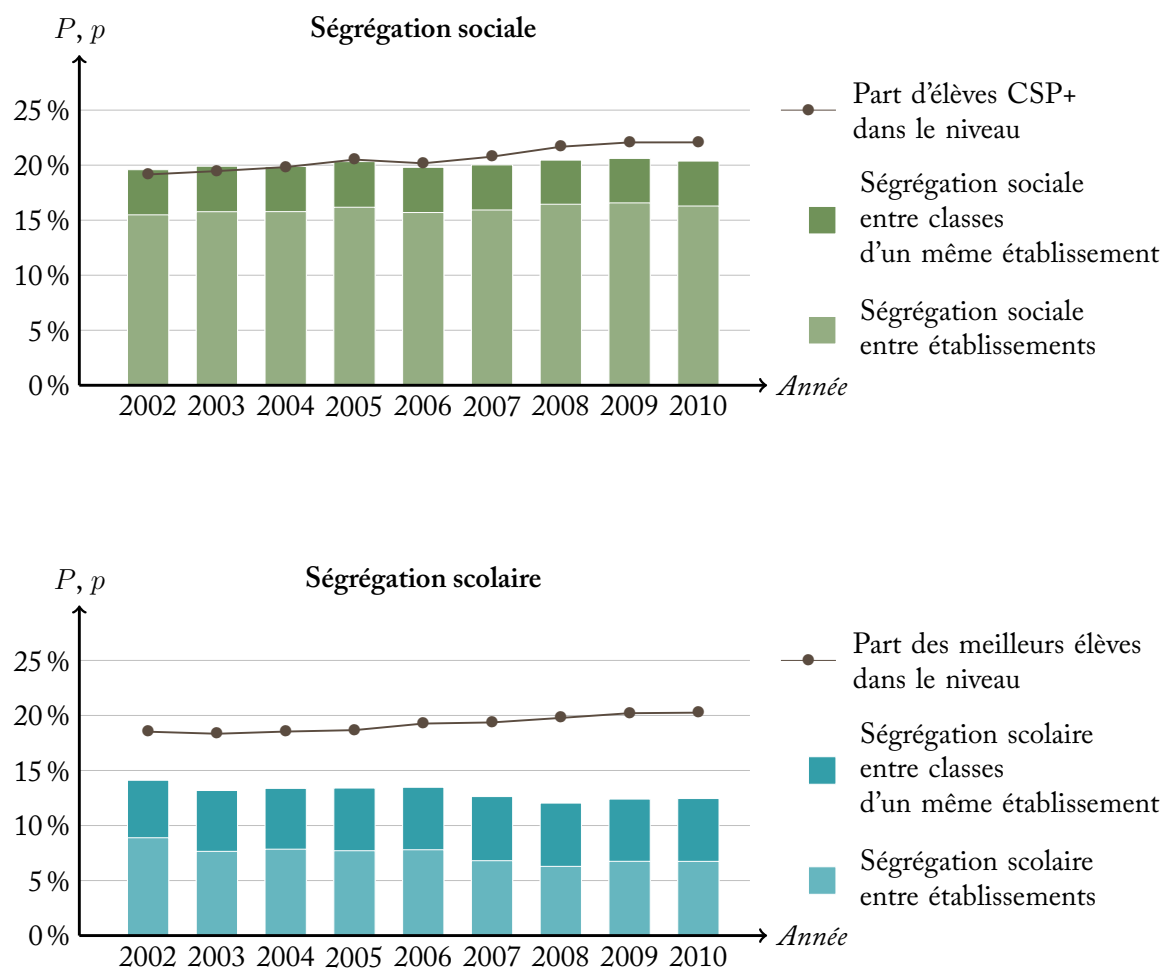


FIGURE 3.10 – Évolution de la ségrégation sociale et scolaire au niveau national en sixième entre 2002 et 2010, mesurée par l'indice d'exposition.

Lecture : en 2006, la ségrégation scolaire en sixième, mesurée par l'indice d'exposition, était de 7,8% entre établissements, et 13,5 % entre classes. La part des meilleurs élèves y était de 19,3 %.

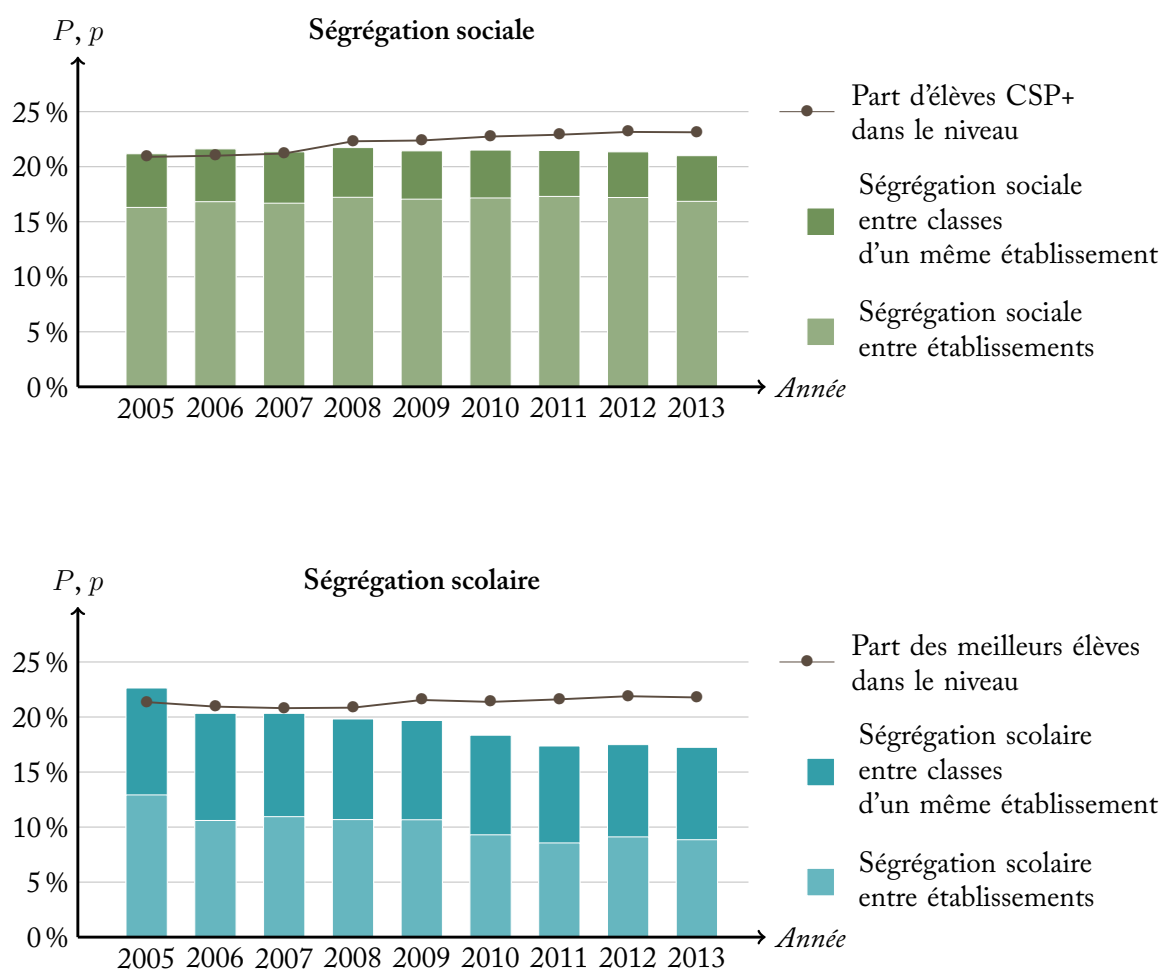


FIGURE 3.11 – Évolution de la ségrégation sociale et scolaire au niveau national en troisième entre 2005 et 2013, mesurée par l'indice d'exposition.

Lecture : en 2009, la ségrégation sociale en troisième, mesurée par l'indice d'exposition, était de 17,1% entre établissements, et 21,4% entre classes. La part des meilleurs élèves y était de 22,4%.

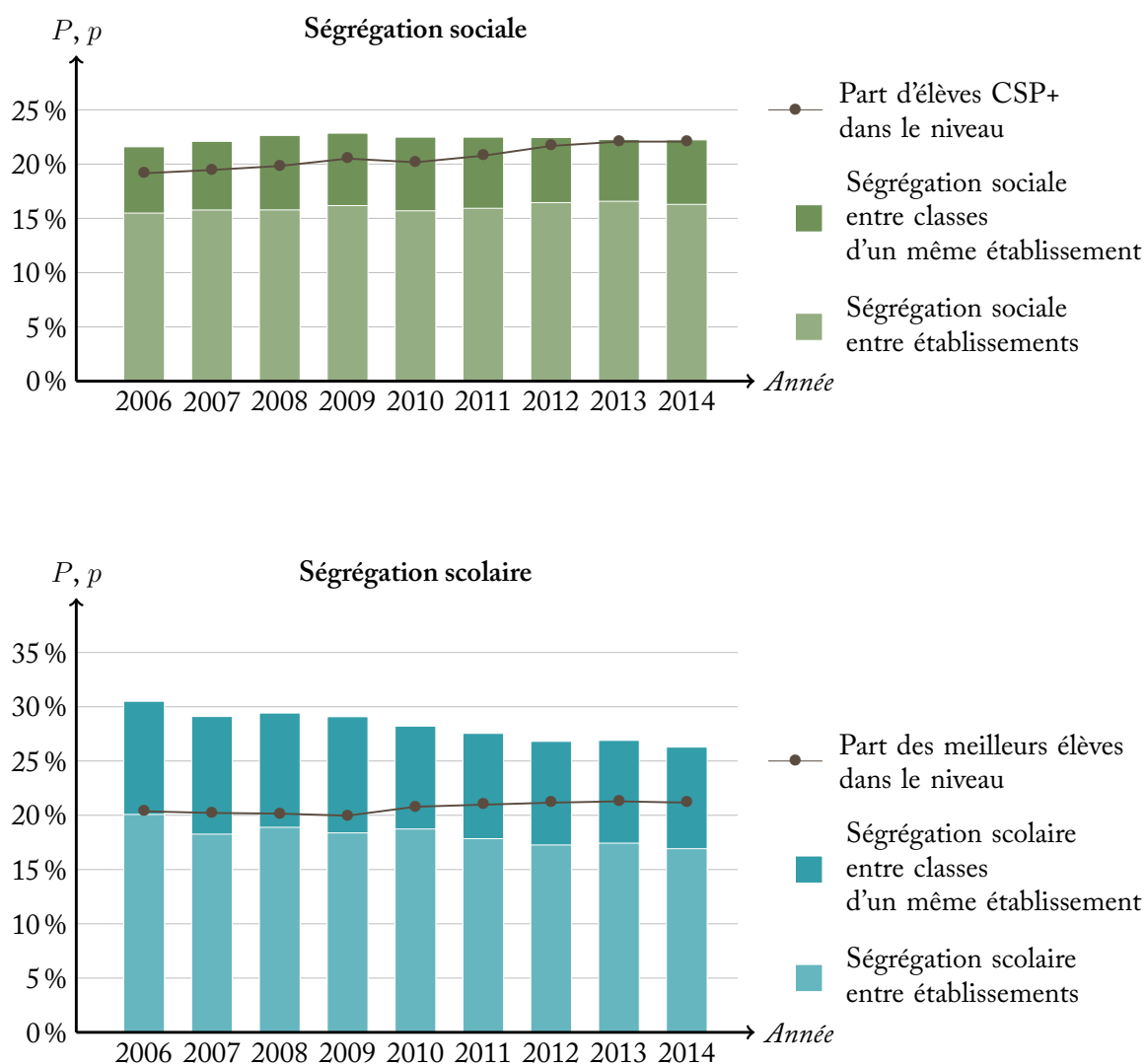


FIGURE 3.12 – Évolution de la ségrégation sociale et scolaire au niveau national en seconde entre 2006 et 2014, mesurée par l'indice d'exposition.

Lecture : en 2010, la ségrégation scolaire en seconde, mesurée par l'indice d'exposition, était de 18,7 % entre établissements, et 28,2 % entre classes. La part des meilleurs élèves y était de 20,8 %.

Nos résultats indiquent tout d'abord une stabilité de la ségrégation sociale au niveau national. Les indices de ségrégation sociale entre établissements et au sein de ceux-ci varient très peu d'une année sur l'autre ; les variations annuelles sont comparables à l'erreur-type des indices et il ne se dégage pas de tendance à la hausse ou à la baisse. Il faut noter que cette stabilité ne concerne que les chiffres agrégés au niveau national. Suite à l'assouplissement de la carte scolaire en 2007, plusieurs études ont observé une hausse de la ségrégation entre établissements dans certaines agglomérations (Merle, 2010 ; Ben Ayed et al., 2013). Ces hausses locales sont compensées par des diminutions de la ségrégation sur d'autres territoires, ce qui a pour effet de maintenir à un niveau stable la valeur de la ségrégation au niveau national, un résultat conforme à ceux de Fack and Grenet (2012).

En ce qui concerne la ségrégation scolaire, nous observons une diminution sensible et régulière de l'indice d'exposition au cours de la période étudiée. À l'exception de la terminale, cette baisse a lieu à tous les niveaux de l'enseignement secondaire, et elle est particulièrement prononcée au collège, où l'indice de ségrégation diminue de 2 à 3 % par an. Cette baisse s'explique à la fois par une diminution de la ségrégation entre établissements et de la ségrégation entre classes d'un même établissement, avec cependant une contribution plus forte de la dimension inter-établissements. Là encore, cette diminution de la ségrégation au niveau national masque des disparités locales : à Paris par exemple, Fack et al. (2014) ont observé une hausse de la ségrégation scolaire après 2008, notamment en raison du poids important donné aux notes dans le barème d'Affelnet, le logiciel d'affectation des élèves à l'entrée au lycée.

Les explications possibles à cette baisse sont multiples : assouplissement de la carte scolaire, procédures d'affectation des élèves, modification de l'offre scolaire et de l'allocation des moyens aux établissements, etc. Ces explications peuvent se combiner différemment en fonction des territoires, et il sera nécessaire de procéder à des analyses plus fines, au niveau local, pour comprendre leurs rôles respectifs, ce qui est hors du cadre de cette étude.

Ainsi, nous constatons au niveau national une diminution tendancielle de la ségrégation scolaire à tous les niveaux de la scolarité, qui ne s'accompagne pas d'une baisse de la ségrégation (*bootstrap*), mesuré l'indice de ségrégation pour chaque ré-échantillonnage et calculé l'écart-type de cette distribution.

sociale. Cette diminution au niveau national masque de fortes disparités locales qui doivent être analysées plus finement.

Conclusion

Cette étude dresse un panorama complet de la ségrégation sociale et scolaire entre les établissements scolaires français et entre les classes de ces établissements, de la sixième à la terminale. Elle est la première à apporter des éléments quantitatifs sur toutes ces dimensions. Elle fait le constat d'une ségrégation sociale importante qui est en grande partie le reflet de la ségrégation résidentielle, au collège comme au lycée. La ségrégation scolaire, deux fois plus forte au lycée qu'au collège, s'explique autant par la ségrégation résidentielle que par les inégalités générées par les compositions de classes.

Alors qu'en l'absence de ségrégation, chaque collégien compterait 22 % d'élèves CSP+ dans son collège, les élèves eux-même CSP+ en comptent 34 % et les autres seulement 18 %. De même, si on considère les meilleurs élèves (tels que repérés par le diplôme national du brevet), ils représenteraient 22 % de chaque classe en l'absence de ségrégation au collège : en réalité, les meilleurs élèves en comptent en moyenne 36 % dans leur classe, et les autres seulement 18 %. Cet écart de 18 points double en classe de première, une augmentation qui s'explique essentiellement par l'orientation.

Ces moyennes nationales cachent d'importantes disparités. On compte tout d'abord une minorité d'établissements dont les classes aisées sont presque absentes (5 % des collèges comptent 3 % ou moins d'élèves CSP+) et à l'inverse 5 % d'établissements privilégiés où leur proportion dépasse le triple de la moyenne nationale. Un constat équivalent est fait en termes de ségrégation scolaire.

La ségrégation entre établissements est un phénomène prédominant en zones urbaines. En effet, les établissements des zones rurales recrutent sur un rayon plus important, ce qui contribue à créer de la mixité sociale et scolaire. La ségrégation entre classes au sein des établissements est, elle, un phénomène plus universel : on observe très peu de variations de son intensité entre

académies, à l'exception notable des académies ultramarines où elle prend des valeurs très élevées.

Les explications de la ségrégation entre classes d'un même établissement sont variées. Il faut tout d'abord noter qu'au collège, 75 % de la ségrégation sociale et 55 % de la ségrégation scolaire s'explique simplement par le hasard : constituer les classes aléatoirement mène en effet à des niveaux de ségrégation comparables à ceux observés. À l'inverse, on identifie donc 25 % à 45 % d'établissements dont la politique de composition des classes ségrège activement les élèves. Cette ségrégation « active » est en partie le résultat de l'affectation des élèves à leurs classes en fonction de leurs options, comme le parcours bilangues en sixième-cinquième, le latin à partir de la cinquième ou la section européenne à partir de la quatrième. Ces options ne sont pas proposées de manière homogène dans tous les territoires, et donnent lieu à un regroupement partiel des élèves par classes au sein des établissements, ce qui génère une hiérarchie sociale et scolaire entre ces classes.

Les facteurs de ségrégation sociale et scolaire sont multiples et complexes. Nous avons pointé ici certains des paramètres qui la structurent. Il ne suffirait cependant pas de s'attaquer à ces paramètres structurels pour effacer durablement la ségrégation, qui est une conséquence inévitable d'un système scolaire compétitif.

3.6 Bibliographie

- Baluteau, F. (2013). *Enseignements au collège et ségrégation sociale*. Académia, Louvain-la-Neuve (Belgique).
- Bell, W. (1954). A probability model for the measurement of ecological segregation. *Social Forces*, 32 :357–364.
- Ben Ayed, C., Broccolichi, S., and Monfroy, B. (2013). Quels impacts de l'assouplissement de la carte scolaire sur la ségrégation sociale au collège ? tendances nationales et déclinaisons locales contrastées. *Éducation & formations*, (83) :39–57.

- Carrington, W. J. and Troske, K. R. (1997). On measuring segregation in samples with small units. *The Journal of Business and Economic Statistics*, 15(4) :402–409.
- Coleman, J., Hoffer, T., and Kilgore, S. (1982). Achievement and segregation in secondary schools : A further look at public and private school differences. *Sociology of Education*, 55(2) :162–182.
- Duncan, O. D. and Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, 20(2) :210–217.
- Duru-Bellat, M. and Mingat, A. (1997). La constitution de classes de niveau dans les collèges ; les effets pervers d’une pratique à visée égalisatrice. *Revue française de sociologie*, 38(4) :759–789.
- Fack, G. and Grenet, J. (2012). Rapport d’évaluation de l’assouplissement de la carte scolaire. Technical report, École d’économie de Paris et Cepremap.
- Fack, G., Grenet, J., and Benhenda, A. (2014). L’impact des procédures de sectorisation et d’affectation sur la mixité sociale et scolaire dans les lycées d’île-de-france. Technical Report 3, Institut des politiques publiques.
- Felouzis, G. (2003). La ségrégation ethnique au collège et ses conséquences. *Revue française de sociologie*, 44(3) :413–447.
- Frankel, D. M. and Volij, O. (2011). Measuring school segregation. *Journal of Economic Theory*, 146(1) :1–38.
- Le Donné, N. and Rocher, T. (2010). Une meilleure mesure du contexte socio-éducatif des élèves et des écoles. *Éducation & formations*, 79 :103–115.
- Ly, S. T., Maurin, É., and Riegert, A. (2014). La mixité sociale et scolaire en île-de-france : le rôle des établissements. Technical Report 4, Institut des politiques publiques.
- Ly, S. T. and Riegert, A. (2015). Measuring social environment mobility. Document de travail G2015/04, Institut national de la statistique et des études économiques.
- Massey, D. S. and Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces*, 67(2) :281–315.

- Maurin, É. (2004). *Le ghetto français : enquête sur le séparatisme social*. La république des idées. Seuil.
- Merle, P. (2010). Structure et dynamique de la ségrégation sociale dans les collèges parisiens. *Revue française de pédagogie*, 1(170) :73–85.
- Monso, O., Givord, P., Guillerm, M., Murat, F., and Afsa, C. (2015). Quels outils pour mesurer la ségrégation dans le système éducatif? une application à la composition sociale des collèges français. Présentation aux Journées de Méthodologie Statistique.
- Rathelot, R. (2012). Measuring segregation when units are small : A parametric approach. *Journal of Business & Economic Statistics*, 30(4) :546–553.
- Reardon, S. F. and O’Sullivan, D. (2004). Measures of spatial segregation. *Sociological methodology*, 34(1) :121–162.
- Thaurel-Richard, M. and Murat, F. (2013). Évolution des caractéristiques des collèges durant la mise en œuvre de l’assouplissement de la carte scolaire de 2007. *Éducation & formations*, (83) :11–23.
- Theil, H. (1972). *Statistical Decomposition Analysis*. North Holland.

3.A Les indices de ségrégation

3.A.1 Définitions

La littérature en économie et sociologie utilise plusieurs indices de ségrégation. Nous rappelons ici la définition des trois indices les plus fréquemment utilisés. Nous en donnons les définitions mathématiques ainsi que les interprétations simples, lorsqu’elles existent.

Nous introduisons tout d’abord les notations suivantes :

- On note N la taille de la population et on note les élèves $i \in \{1, \dots, N\}$.
- On considère un découpage de la population en deux groupes notés A et B. Le groupe A est le groupe dit « de référence » et le groupe B contient le reste de la population. On appelle p la proportion du groupe A dans la population.

- On considère par ailleurs un découpage en « environnements fréquentés » à K unités (par exemple, K écoles ou K classes), chaque unité étant repéré par la lettre $k \in \{1, \dots, K\}$.
- On note N_k la taille de l'unité k et p_k la proportion d'élèves du groupe A dans cette unité.
- On note μ_i l'exposition de l'élève i au groupe A, c'est-à-dire la proportion d'élèves du groupe A dans l'unité fréquentée par l'élève i . Ainsi, si l'élève i appartient à l'unité k , $\mu_i = p_k$.

Selon ces notations, la moyenne des p_k pondérée par les tailles N_k et la moyenne de μ_i sur la population sont toutes deux égales à p : l'exposition moyenne au groupe de référence est égale à la part de ce groupe dans la population.

Définition 5 (Indice d'exposition) *L'indice d'exposition, défini par Bell (1954), a la formule suivante :*

$$P = \frac{1}{p(1-p)} \sum_{k=1}^K \frac{N_k}{N} (p_k - p)^2 = \frac{1}{p(1-p)} \cdot \frac{1}{N} \sum_{i=1}^N (\mu_i - p)^2 \quad (3.1)$$

La définition est proche de celle de l'indice de dissimilarité, en utilisant des écarts quadratiques au lieu des écarts absolus. Il peut être interprété de deux manières.

- Tout d'abord, on peut le voir comme une *part de variance expliquée* (un « R^2 ») par les découpages des environnements fréquentés : il indique dans quelle mesure la variance de la variable d'appartenance au groupe de référence est expliquée par les classes ou les écoles. En ce sens, il a l'avantage d'être décomposable : lorsqu'on calcule la ségrégation entre classes, le résultat peut être écrit comme la somme d'un terme inter-établissements et d'un terme intra-établissement, ce qui sera très utile dans notre analyse. Le seul autre indice vérifiant cette propriété est l'indice d'entropie H , qui a cependant le défaut de ne pas avoir d'interprétation simple.
- On peut également voir l'indice d'exposition P comme *l'écart d'exposition au groupe A entre le groupe A et le groupe B*. En effet, si on note μ^A l'exposition moyenne des élèves du groupe A aux élèves de ce même groupe, et μ^B l'exposition des autres élèves (groupe B) au groupe A, on peut alors écrire :

$$P = \mu^A - \mu^B \quad (3.2)$$

Définition 6 (Indice de dissimilarité) *L'indice de dissimilarité, introduit par Duncan and Dun-*

can (1955), est défini par la formule :

$$D = \frac{1}{2p(1-p)} \sum_{k=1}^K \frac{N_k}{N} |p_k - p| = \frac{1}{2p(1-p)} \cdot \frac{1}{N} \sum_{i=1}^N |\mu_i - p| \quad (3.3)$$

Il s'agit de l'écart absolu moyen entre l'exposition d'un élève au groupe de référence et la part de ce groupe dans la population, divisée par la valeur maximale que peut prendre cet écart moyen, atteinte en situation de ségrégation complète. Cette normalisation permet de faire en sorte que D évolue entre zéro et un. L'indice D représente également la part minimale de la population qu'il faudrait déplacer pour atteindre l'équilibre parfait, là encore divisée par la part maximale qu'il faudrait déplacer en cas de ségrégation complète.

Définition 7 (Indice d'entropie, ou indice de Theil) *L'indice d'entropie a été défini par Theil (1972). Il porte parfois le nom de son auteur. Sa définition est la suivante :*

$$H = 1 - \frac{1}{h(p)} \sum_{k=1}^K \frac{N_k}{N} h(p_k) = 1 - \frac{1}{h(p)} \cdot \frac{1}{N} \sum_{i=1}^N h(\mu_i) \quad (3.4)$$

où $h(\cdot)$ est la fonction d'entropie, définie par :

$$h(u) = u \cdot \log_2 \left(\frac{1}{u} \right) + (1 - u) \cdot \log_2 \left(\frac{1}{1 - u} \right) \quad (3.5)$$

Cet indice de ségrégation a des propriétés mathématiques intéressantes et peut être, plus naturellement que les autres, adapté à une analyse à plusieurs groupes. Pour ces raisons, il est de plus en plus populaire chez les économistes ; c'est notamment l'indicateur de choix d'une étude conjointe de l'Insee et de la Depp (Monso et al., 2015). Il n'a cependant pas d'interprétation simple.

3.A.2 Discussion

Chacun de ces indices possède des propriétés mathématiques différentes, et le choix de l'indice dépend du contexte d'utilisation. Nous avons choisi l'indice d'exposition car il est décomposable en une somme de contributions à la ségrégation et qu'il possède une interprétation simple.

L'indice d'entropie est celui qui peut le plus naturellement être utilisé pour mesurer la ségrégation entre plus de deux groupes²⁰.

L'indice de dissimilarité est parfois utilisé au motif qu'il est « insensible à la composition sociale » et qu'il permet donc les comparaisons d'indices entre contextes différents (entre régions, villes, etc.). Selon nous, cet argument ne tient pas. Comme le notent [Coleman et al. \(1982\)](#), la propriété dite d'insensibilité à la composition sociale vérifiée par l'indice de dissimilarité a une signification bien précise qui ne permet pas de justifier des comparaisons entre des territoires ayant des compositions sociales différentes.

En effet, l'indice de dissimilarité est invariant lorsque la proportion du groupe de référence est multiplié par un même facteur dans toutes les unités du territoire étudié. Or, si la proportion de CSP+, par exemple, est multiplié par 1,1 (hausse de 10 %) partout, la proportion de non-CSP+ ne diminue pas systématiquement de la même proportion. En effet, dans un établissement qui passe de 20 à 22 % de CSP+, la proportion de CSP- passera de 80 à 78 %, soit une baisse de 2,5 % ; dans un établissement qui passe de 60 à 66 % de CSP+, la proportion de CSP- passera de 40 à 34 %, soit une baisse de 15 %. Cela représente des variations très différentes en terme de vécu des élèves et d'interactions potentielles. Si cette propriété peut avoir un intérêt mathématique, elle ne suffit pas pour comparer les indices de ségrégation de territoires ayant des compositions sociales différentes.

De manière générale, il faut être vigilant dans les comparaisons d'indice et ne pas interpréter de manière absolue des différences de ségrégation entre des territoires ayant des compositions sociales très différentes²¹, et ce quel que soit l'indice de ségrégation choisi. Il convient par ailleurs de toujours accompagner les indices de ségrégation d'une indication de la composition sociale de l'échantillon étudié (par exemple, la proportion du groupe dont on mesure l'isolement dans

20. La séparation des CSP par la Depp en quatre groupes aurait pu justifier l'utilisation de l'indice d'entropie. Nous n'avons pas fait ce choix, car l'indice de ségrégation ainsi obtenu regroupe en un seul chiffre des phénomènes différents. Par exemple, [Monso et al. \(2015\)](#) montrent qu'au cours des dix dernières années à Paris, la ségrégation sociale mesurée par l'indice d'entropie est restée stable, mais que la ségrégation des CSP- a augmenté alors que la ségrégation des autres groupes a légèrement diminué. Nous avons préféré ici montrer séparément les valeurs de la ségrégation des CSP+ et des CSP-.

21. Selon [Reardon and O'Sullivan \(2004\)](#), des faibles variations de la composition sociale permettent d'effectuer des comparaisons sans introduire de biais important.

l'échantillon).

3.B Mesures alternatives de la ségrégation

Nous avons choisi dans cette étude de mesurer la ségrégation en utilisant l'indice d'exposition et en prenant comme groupes de références les élèves les « favorisés », c'est-à-dire les CSP+ pour mesurer la ségrégation sociale et les meilleurs élèves au DNB pour mesurer la ségrégation scolaire. Dans cette annexe, nous donnons les valeurs de la ségrégation mesurée selon d'autres indicateurs, pour la cohorte d'élèves entrant en sixième en 2007.

- La **FIGURE 3.B.1** donne la valeur de l'indice d'exposition en prenant comme groupes de références les élèves CSP- pour la ségrégation sociale et les élèves en retard scolaire pour la ségrégation scolaire.
- La **FIGURE 3.B.2** donne la valeur de l'indice de dissimilarité D (en utilisant à nouveau les groupes de référence utilisés dans le reste de l'étude). Cet indice, souvent utilisé notamment dans la littérature sociologique, est égal à la part de la population qui devrait changer d'établissement ou de classe pour atteindre un équilibre parfait, divisée par ce que serait cette part en cas de ségrégation complète. Contrairement à l'indice d'exposition, celui-ci ne peut pas être décomposé en une composante inter-établissements et une composante intra-établissement : nous calculons séparément la ségrégation entre établissements puis la ségrégation entre classes, mais la différence entre les deux ne peut pas être interprétée comme une valeur de la ségrégation intra-établissement. C'est la raison pour laquelle nous n'utilisons pas de barres empilées dans cet histogramme.
- La **FIGURE 3.B.3** représente la valeur de l'indice de Theil H . Cet indice a une construction mathématique très proche de l'indice d'exposition que nous utilisons principalement dans l'étude ; il n'a cependant pas d'interprétation simple. Les valeurs obtenues sont également proches des valeurs de l'indice d'exposition.

Ces mesures alternatives de la ségrégation sont donnés à titre de comparaison avec d'autres études. Nous n'entrons pas plus en détail dans leur analyse, nous notons simplement que ces chiffres ne mènent pas à un diagnostic différent de celui effectué dans cette étude.

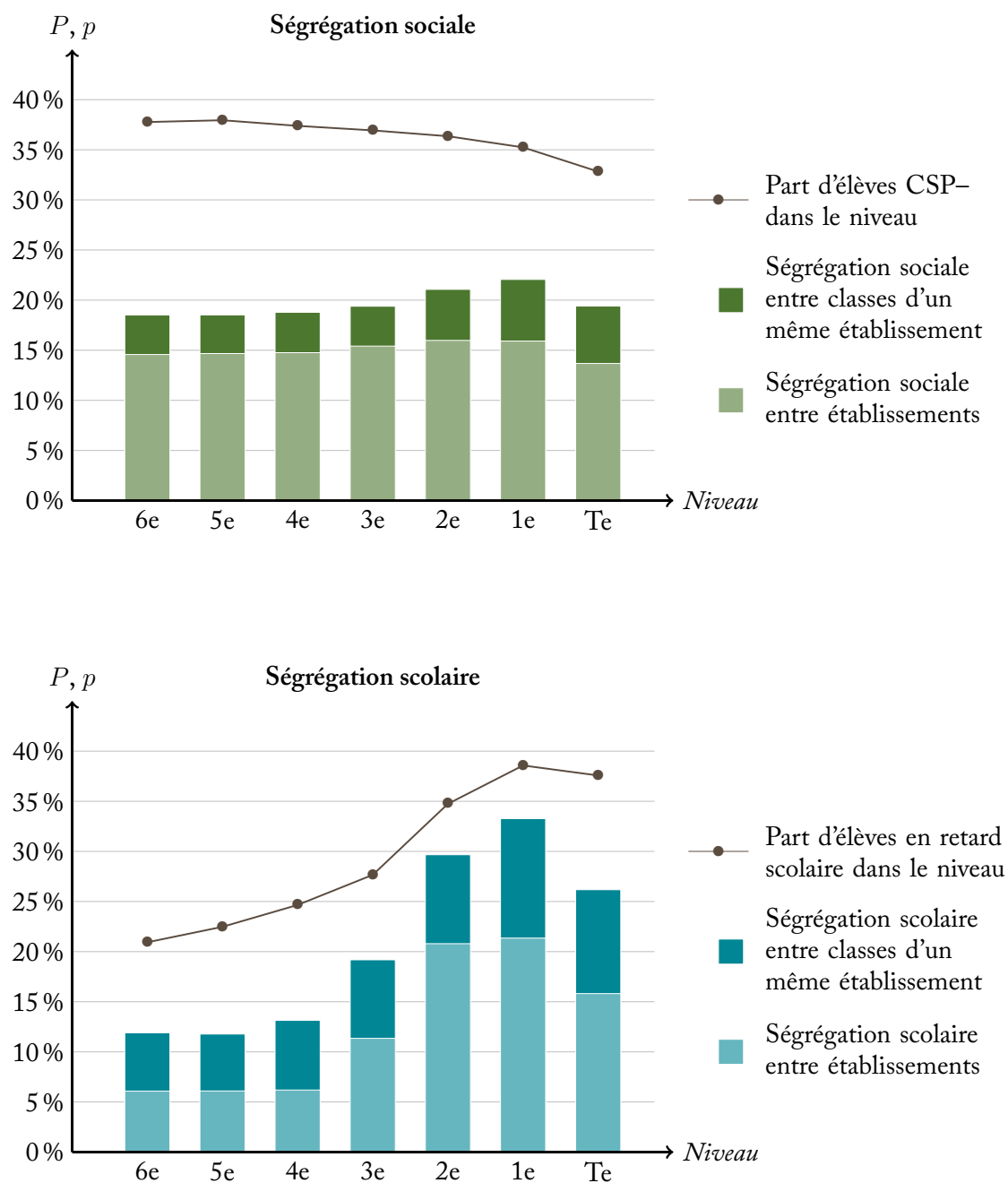


FIGURE 3.B.1 – Ségrégation sociale et scolaire au niveau national, mesurée par l'indice d'exposition à chaque niveau pour la cohorte 2007, en utilisant les élèves « défavorisés » comme groupe de référence.

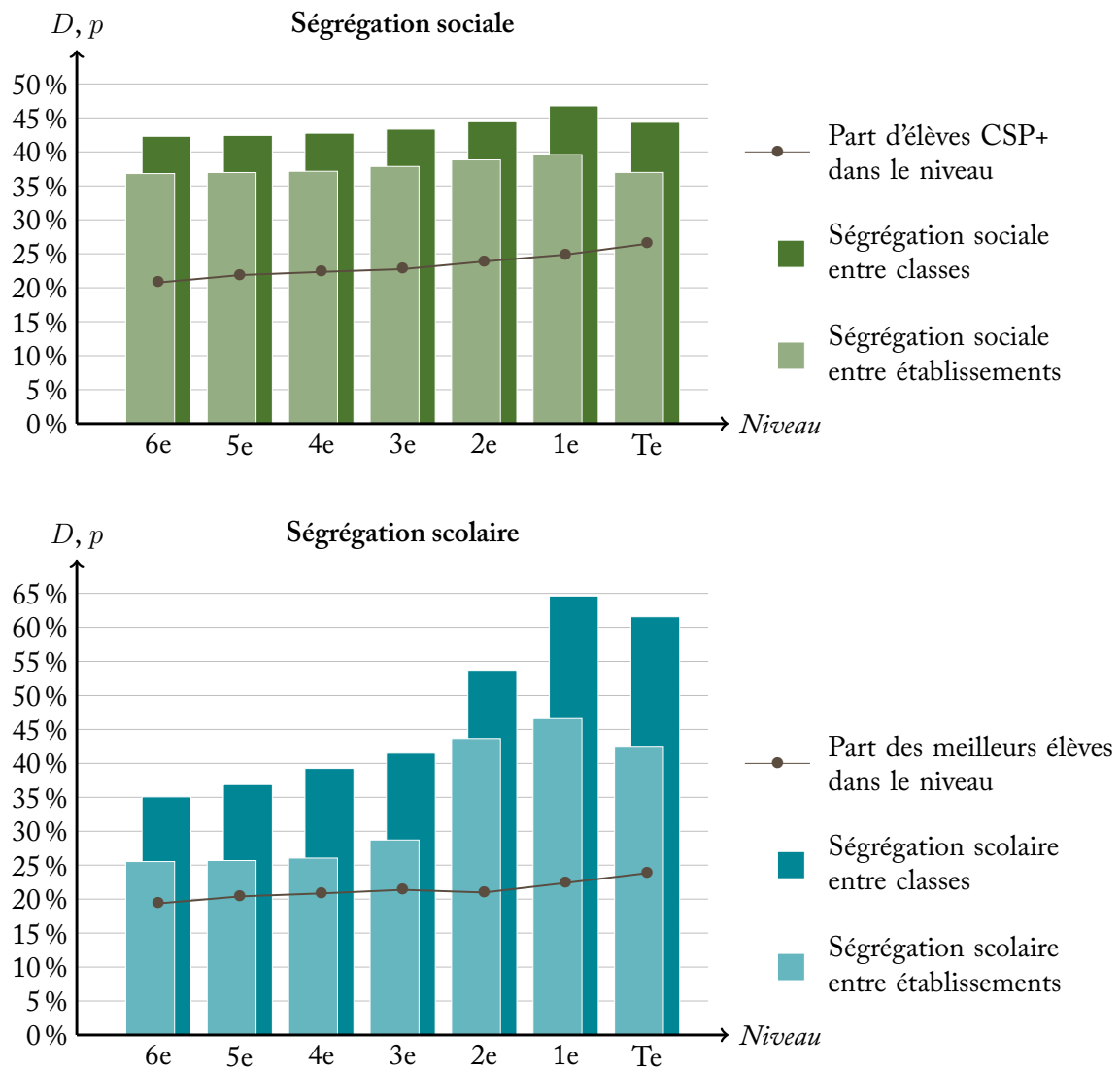


FIGURE 3.B.2 – Ségrégation sociale et scolaire au niveau national, mesurée par l'indice de dissimilarité à chaque niveau pour la cohorte 2007.

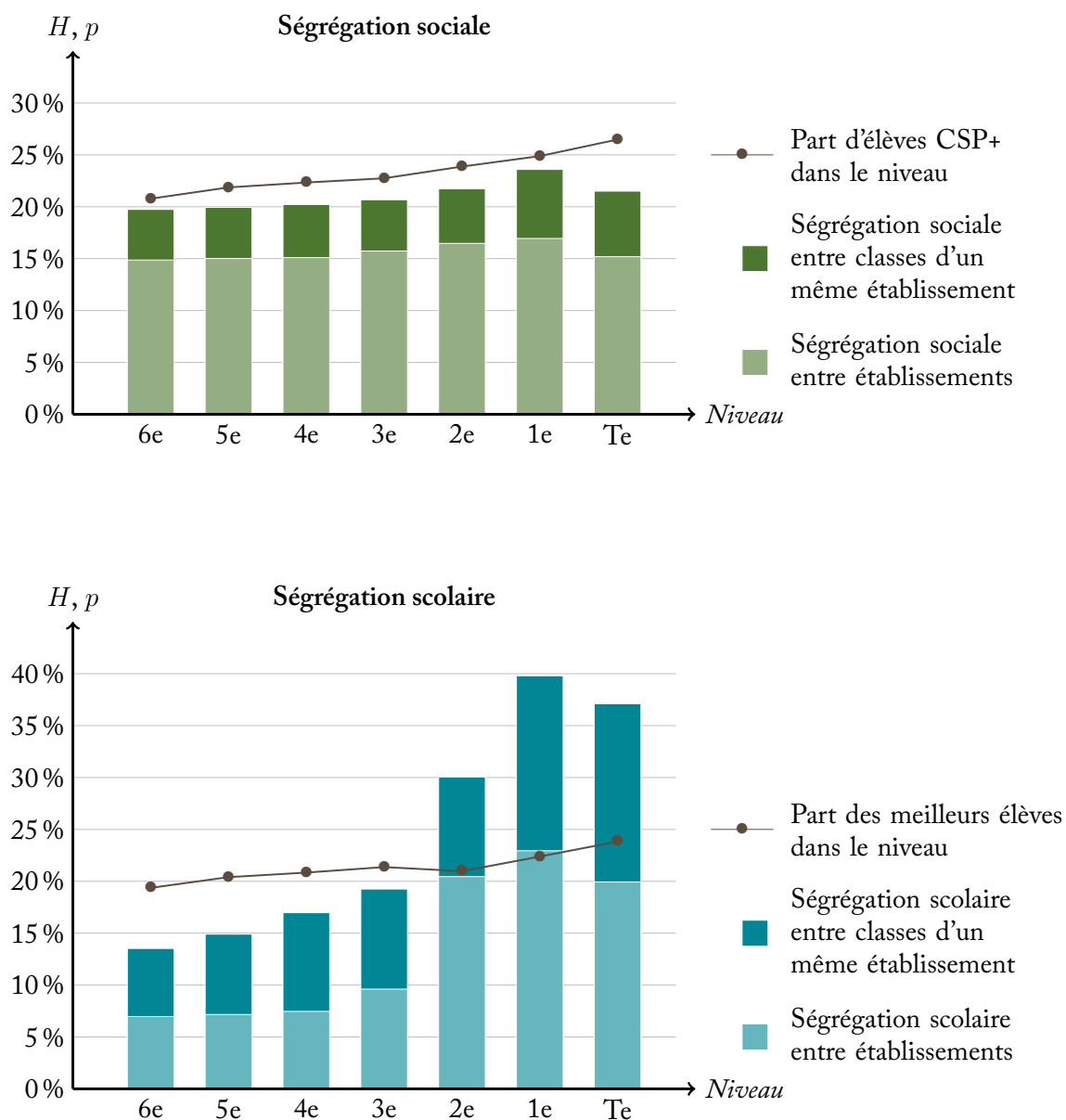


FIGURE 3.B.3 – Ségrégation sociale et scolaire au niveau national, mesurée par l'indice de Theil à chaque niveau pour la cohorte 2007.

Chapitre 4

Measuring Social Environment Mobility

Ce chapitre a été écrit avec Son Thierry Lx.

Il a été publié en tant que document de travail du Département des études économiques de l'Insee (numéro G2015/04).

Les auteurs souhaitent remercier Luc Behaghel, Sandra Black, Laurent Gobillon, Malik Koubi, Éric Maurin, Magne Mogstad, Fabrice Murat ainsi que les participants au lunch séminaire d'économie appliquée de l'École d'économie de Paris et au séminaire du département des études économiques de l'Insee pour leurs commentaires, questions et remarques qui ont permis de faire progresser ce document de travail.

4.1 Introduction

School or residential segregation between social or ethnic groups has been a great concern for social scientists and policymakers for many decades. By depriving minorities from key resources

in their environment, the spatial separation of individuals by social background or ethnicity may hamper their opportunities and thus reinforce inequalities (Crain and Strauss, 1985; Wells and Crain, 1994; Cutler and Glaeser, 1997). Segregation may also be viewed as a threat to democracy by developing uneasiness about interracial contacts, if not racial hostility (Braddock II and McPartland, 1989). This is suggested in particular in recent works on social networks, which show that geographical proximity could strongly reduce the homophilic feature of interactions between individuals (Currarini et al., 2009; Marmaros and Sacerdote, 2006). For example, Camargo et al. (2010) report that different-race students are as likely to become friends as same-race students once they have been (randomly) assigned as roommates (also see Marmaros and Sacerdote, 2006). By depriving individuals from diversified interactions at an early stage of development, school segregation may therefore have long run implications on relationships between social or racial groups¹.

However, the high interest of scholars on segregation issues has stalled at the conceptual and technical challenges raised by its measurement. A rich literature has spawned from this issue, with recent papers (Echenique and Fryer, 2007; Frankel and Volij, 2011) that added new measures and concepts to the existing battery of segregation indices (Duncan and Duncan, 1955; Massey and Denton, 1988). Interestingly enough, all these papers share a common limited feature: they only consider segregation in a cross-sectional perspective. In other words, all segregation indices that have been developed only measure how much individuals from different groups are spatially separated from each other *at a given point of time*. Yet the scope of the conclusions that can be drawn from such indices is somewhat restricted: if individuals move a lot from an environment to another across years, segregation in a longitudinal perspective may be substantially lower than it appears from a cross-sectional point of view. For instance, cross-sectional indices may inform us about the extent to which students in a school are segregated between classes with mostly white students on the one hand, and classes with mostly black students on the other hand. They do not tell us whether each student stays every year in the same kind of classroom, or if instead, they tend to move between "rich" and "poor" classes every year. The implications for students' concrete experiences of social isolation is very different depending on

1. In this direction, Camargo et al. (2010) also report interesting long run effects of diversity on white students. According to the authors, they are more likely to have black friends in the future when they are assigned to a black roommate in college.

the scenario.

This paper proposes to fill this gap in the literature by providing the first theoretical and empirical analysis of social environment mobility and its consequences on segregation. In doing so, our work closely relates to the strand of literature that embeds earnings mobility in the measurement of income inequality, following [Shorrocks \(1978b,a\)](#) in particular. By examining inequalities of permanent or lifetime income instead of "snapshot" income, this literature has clearly established the equalizing force of mobility (see *e.g.* [Bonhomme and Robin, 2009](#), in the case of France). For example, [Bowlus and Robin \(2004\)](#) report that lifetime income inequality in the U.S. is 40% less than in instantaneous earnings. Obviously, considering income mobility also yields a different vision of the evolution of income inequality over time ([Moffitt and Gottschalk, 2002](#)) or of cross-country comparisons ([Gottschalk and Spolaore, 2002](#)). In the same spirit, *social environment mobility* (SEM throughout the paper) may have an equalizing effect on the distribution of environments across individuals.

The paper is organized as follows. In section [4.2](#), we review the literature on the measurement of income mobility and of segregation. We explain why considering income mobility leads to lower indices of income inequality and we present the main tools to measure income mobility. We also present the segregation indices that prevail in the literature on segregation. In section [4.3](#), we introduce the concept of social environment mobility. We first show how segregation indices are related to income inequality indices, in order to adapt the income mobility indices to the case of social environment mobility. We present our main result which is that unlike income mobility, social environment mobility can not reach 100 percent as long as there is some degree of segregation at one point in time. In other words, while lifetime income inequality may be reduced to zero even if instantaneous inequality is not null, there is a lower bound in the case of segregation: it is not possible that all individuals in the population experience the same average exposure to a reference group if that reference group is not evenly distributed geographically at all times. We provide tools that help using our social environment mobility index with real data and we conclude with a case study on French middle schools (section [4.4](#)).

4.2 Literature review

There is an important and still active literature on the measurement of income mobility and the measurement of segregation. Although social environment mobility (SEM) is conceptually different from income mobility, its measurement uses similar mathematical tools. In this section, we present some of the tools we use as a basis to build our SEM indices.

4.2.1 Measuring income inequality

This section gives a quick overview of the most common tools to measure income inequality. We consider a population of N individuals $i \in \{1, \dots, N\}$. Individual i has income y_i . The income distribution has a cumulative distribution function $\mathcal{F}(y)$ and average $\bar{y}$.

Inequality indices are functions of the income distribution $I(\mathcal{F})$ that belong to the unit interval $[0, 1]$. They equal zero when all individuals have the same income and one when all the income is owned by one individual only. Different indices may have different variations between these two extreme cases depending on their axiomatic properties.

The *Lorenz curve*, introduced by **Lorenz (1905)**, is one of the most popular tools to study income inequality. The curve is defined by the following function:

$$\mathcal{L}(u) = \frac{\int_0^u \mathcal{F}^{-1}(t) dt}{\int_0^1 \mathcal{F}^{-1}(t) dt} = \frac{1}{\bar{y}} \int_0^u \mathcal{F}^{-1}(t) dt \quad \forall u \in [0, 1] \quad (4.1)$$

Note that the joint knowledge of $\mathcal{L}$ and $\bar{y}$ is equivalent to knowing the full distribution of income (for instance, $\bar{y}\mathcal{L}'(u) = \mathcal{F}^{-1}(u)$).

The interpretation of the Lorenz curve is the following. $\mathcal{L}(20\%)$ is the share of all the income that is earned by the 20% individuals with the lowest income.² It follows that $\mathcal{L}(u) \leq u$ for all u and the Lorenz curve is below the 45-degree line. In addition, $\mathcal{L}(0) = 0$ and $\mathcal{L}(1) = 1$, so

2. Indeed, $\int_0^u \mathcal{F}^{-1}(t) dt = \int_0^{\mathcal{F}^{-1}(u)} t f(t) dt$ is the total income of the fraction u of the poorest individuals in the population, therefore $\mathcal{L}(u)$ is equal to that income divided by the total income of the whole population.

the curve and the line cross at $(0, 0)$ and $(1, 1)$. If there is no inequality (everybody has the same income), $\mathcal{L}(u) = u$ for all u . At the other end, when all income is owned by one individual, $\mathcal{L}(u) = 0$ if $u < 1$ and $\mathcal{L}(1) = 1$.³ Between these two extremes, the Lorenz curve lies between the x -axis and the 45-degree line. The closer it is to the 45-degree line, the lower inequality is.

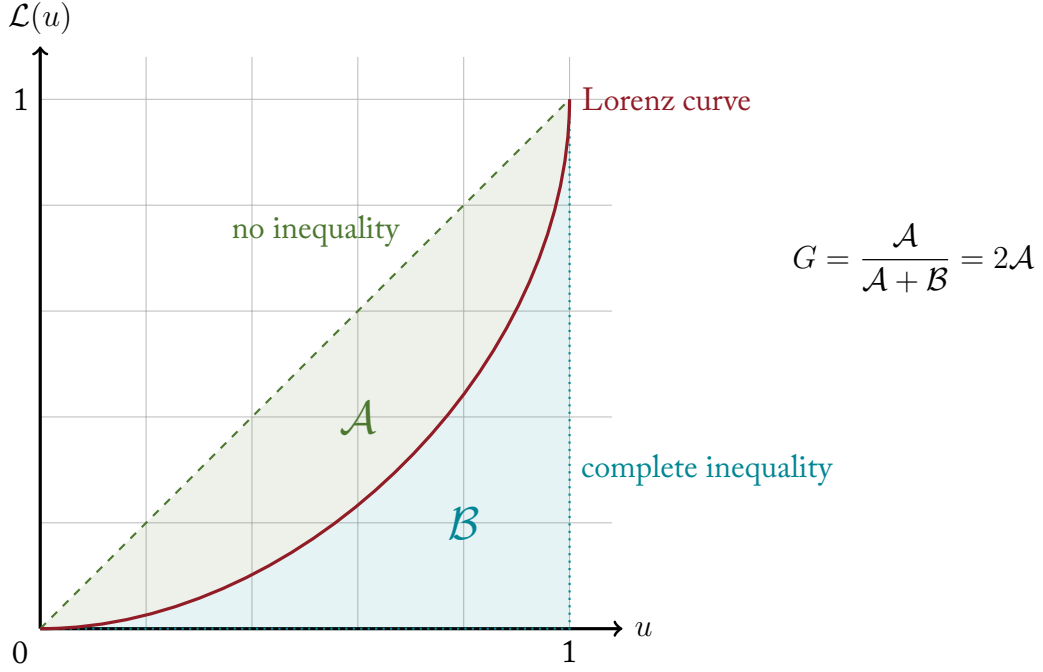


FIGURE 4.1 – The Lorenz curve for income inequality

The most popular inequality index – the Gini index – has a simple geometric relationship with the Lorenz curve. It is equal to the area between the Lorenz curve and the 45-degree line, divided by its highest possible value ($1/2$) so that it belongs to the unit interval. It can be written:

$$G = \frac{\int_0^1 [u - \mathcal{L}(u)] du}{1/2} = 1 - 2 \int_0^1 \mathcal{L}(u) du \quad (4.2)$$

Note that the Gini index is also equal to half the average absolute difference between any two individuals' incomes, *i.e.*

$$G = \frac{1}{2N^2} \frac{1}{y} \sum_{i=1}^N \sum_{j=1}^N |y_i - y_j| \quad (4.3)$$

3. In reality, $\mathcal{L}(u) = 0$ if $u \leq 1 - 1/N$, but we assume that $N \gg 1$.

See appendix 4.A.1 for a proof of the equivalence of definitions (4.2) and (4.3).

Remark: The Gini index belongs to a broader class of inequality measures which have the following form (Yaari, 1988; Aaberge, 2001):

$$J_p(\mathcal{L}) = 1 - \int_0^1 p(u) d\mathcal{L}(u) \quad (4.4)$$

where the weighting function p is a non-increasing function with unit integral and $p(1) = 0$. The choice of p belongs to the social planner and depends on the desired axiomatic properties of the inequality index. The Gini index is associated with the weighting function $p_G(u) = 2(1 - u)$. Indeed, if we integrate by parts:

$$\begin{aligned} J_{p_G}(\mathcal{L}) &= 1 - \int_0^1 2(1 - u) d\mathcal{L}(u) \\ &= 1 - [2(1 - u)\mathcal{L}(u)]_0^1 + \int_0^1 (-2)\mathcal{L}(u) du \\ &= 1 - 2 \int_0^1 \mathcal{L}(u) du \\ &= G \end{aligned}$$

4.2.2 The equalizing force of income mobility

The purpose of income mobility indices is to show that instantaneous income inequality indices do not tell the full story of the income distribution. Income inequality is the combination of two very different phenomena at work. First, people earn different incomes throughout their lives (for instance their incomes increase with seniority or decrease when they retire), which generates inequality when looking at the whole population at one point in time. Second, within age groups, a number of personal characteristics affect income (skills, education, etc.). These two contributions should be analyzed separately: different societies may have different standards as to what level of inequality of each type is acceptable.

Income mobility indices measure to what extent people are likely to move in the income distribution. There are essentially two approaches to measuring this phenomenon. The first approach consists in measuring correlations between current and past rankings in the income

distribution. The second approach consists in comparing the distribution of instantaneous incomes to the distribution of lifetime incomes.

Consider the following example with two individuals in two periods of time, and two scenarios. In scenario A, the income distribution is $(y_1 = 1,000; y_2 = 2,000)$ in both periods. In scenario B, the income distribution is $(y_1 = 1,000; y_2 = 2,000)$ in period 1 and incomes are swapped in period 2: $(y_1 = 2,000; y_2 = 1,000)$. In both scenarios and in both periods, any instantaneous inequality index would have the exact same value. For instance, the Gini index would equal $G = 1/6$. These two scenarios could thus not be distinguished by analyzing instantaneous inequality indices solely. Yet scenario B exhibits more income mobility than scenario A, as both individuals exchange places in the income distribution: therefore, the correlation between current and past rankings is lower, and the inequality in lifetime incomes is also lower (zero if we consider that the interest rate is zero).

High income mobility is a signal of equal opportunities: if people are likely to shift in the distribution, this means that people at the bottom of the distribution may make their way up to the top, and conversely people who are initially in the top of the distribution do not automatically keep their privileges.

If there is no income mobility at all, *i.e.* if all individuals keep the same (relative) income throughout their life, individuals also keep their positions in the income distribution, and the lifetime income is proportional to any instantaneous income. Therefore, the *perfect income rigidity* scenario leads to perfect correlation between present and past positions in the distribution, and identical measures of inequality based on instantaneous of lifetime income.

At the opposite end, defining *perfect income mobility* is more complicated. One possible definition would be to consider the case where past incomes have no influence on the current income (Prais, 1955). In this case, the correlation between past and current positions in the distribution would be equal to zero, and an index of inequality based on lifetime income would tend towards zero as the number of time periods would grow to infinity.

However, this definition does not describe the highest possible mobility. For instance, we could imagine a case where poor people have better odds being rich in the next period, while rich people will more likely get poor. In this case, the correlation between past and current ranks in the distribution would be negative, and the inequality index based on lifetime income may go to zero very rapidly.

In the rest of this section, we present income mobility indices that are based on the two classes of methods introduced above, *i.e.* measuring the correlation between past and current rankings, and comparing instantaneous and permanent income distributions. In order to compare present and past positions in the income distribution, a simple possibility is to use autocorrelation coefficients of the series of incomes. However, the literature on social mobility uses transition matrices, from which mobility indices can be derived. We also show two indices that are based on the comparison of instantaneous and lifetime income.

All the indices introduced here take values between zero and one. Mobility is equal to zero in the perfect rigidity scenario. The remainder of the axiomatic properties may vary from one index to another and can be found in the referenced papers.

4.2.2.1 Transition matrices

Transition matrices are a popular tool to measure social mobility, *i.e.* to observe how individuals are likely to move between categories of the population at different points in time. These matrices are especially useful when the population is partitioned into categories such as the socioeconomic status, but they may also be used to study mobility within the distribution of a continuous variable such as the income, by using income quantiles as categories.

Consider a population partitioned into n categories and two points in time ($t \in \{1; 2\}$). Denote $c_i^t \in \{1; \dots; n\}$ the category of individual i at time t . The transition matrix is $T = (t_{jk})_{1 \leq j, k \leq n}$ where $t_{jk} = \mathbb{P}(c_i^2 = k | c_i^1 = j)$ is the share of the population that was in category j in the first period ending up in category k in the second period. A transition matrix should be

stochastic, *i.e.* that the coefficients in each row must add up to one:

$$\sum_{k=1}^n t_{jk} = \sum_{k=1}^n \mathbb{P}(c_i^2 = k | c_i^1 = j) = 1 \quad (4.5)$$

The perfect rigidity scenario's transition matrix is the identity matrix. In the perfect mobility scenario, the initial category has no influence on the final category: all rows are identical. [Shorrocks \(1978b\)](#) argued that the perfect mobility scenario should be associated with a mobility index value of one. However, this axiom associated with a monotonicity axiom would lead to values higher than one when mobility is more than perfect. [Shorrocks \(1978b\)](#)'s recommendation is to limit the value of the index to one in those cases, that he considers very unlikely in the context of income mobility.

An example of such an index is the following:

$$M_T = \min \left(\frac{n - \text{Trace}(T)}{n - 1}; 1 \right) \quad (4.6)$$

Indeed, in the case of perfect rigidity, T is the identity matrix and has trace n , therefore $M_T = 0$. If there is perfect mobility, property (4.5) and the fact that the rows are identical leads to the trace being equal to one, therefore $M_T = 1$.

Another difficulty with indices based on transition matrices is that they greatly depend upon the definition of the categories. For instance, using income percentiles rather than deciles would automatically increase the value of the income mobility index, as the diagonal elements would be lower in proportion. The following sections introduce indices that adress these two limitations.

4.2.2.2 Comparing instantaneous and lifetime inequality

A second class of income mobility indices has been developed by [Shorrocks \(1978a\)](#), published just three months after [Shorrocks \(1978b\)](#). The intuition is that when income mobility is high, inequality indices based on lifetime incomes should be lower than those based on instantaneous incomes. Calling $\mathcal{F}_t$ the income cumulative distribution function (cdf) in period t , $\mathcal{F}^*$ the cdf of lifetime income, and $I(\mathcal{F})$ an income inequality index (such as the Gini index), an

income mobility index is defined by:

$$M = 1 - \frac{I(\mathcal{F}^*)}{\frac{1}{T} \sum_{t=1}^T I(\mathcal{F}_t)} \quad (4.7)$$

where I is a strictly convex function of the relative income distribution. This restriction on I ensures that M belongs to the unit interval (Shorrocks, 1978a). Consider the two extreme cases leading to $M = 0$ and $M = 1$.

In the perfect rigidity scenario, the instantaneous and lifetime income distributions are identical (up to a multiplicative constant in income), therefore the inequality indices – that are functions of *relative* incomes only – are equal, and M is equal to zero.

M equals one if and only if inequality in lifetime income $I(\mathcal{F}^*)$ is equal to zero. This may happen in only three scenarios: (i) if there is no inequality at any point in time, (ii) if income mobility is perfect and the number of periods is infinite or (iii) if income mobility is more than perfect so that inequality in a given period are compensated for in a different period. The first scenario is trivial and the second scenario has no empirical application. The third scenario is what Shorrocks (1978a) calls *complete income mobility*, as opposed to perfect income mobility.

4.2.2.3 Income mobility curves

A third method to measure income mobility has recently been introduced by Aaberge and Mogstad (2014), that also relies on the comparison of instantaneous and lifetime income. The authors extend the concept of Lorenz curve and introduce an income mobility curve which is defined as the difference between the Lorenz curve of the actual lifetime income distribution $\mathcal{F}^*$ and the curve based on a hypothetical, reference lifetime income distribution in perfect rigidity scenario, $\mathcal{F}_R^*$. In the reference distribution $\mathcal{F}_R^*$, "the rank of each individual is the same in every period; this distribution can be formed by assigning the lowest income in every period to the poorest individual in the first period, the second lowest to the second poorest, and so on.". Note that this definition does not give a particular role to the distribution in the first period: keeping the rankings constant and equal to the rankings in any period would lead to the exact same distribution.

Call $\mathcal{L}^*$ the Lorenz curve of the lifetime income distribution $\mathcal{F}^*$ and $\mathcal{L}_R^*$ the Lorenz curve of the reference distribution $\mathcal{F}_R^*$. The income mobility curve is defined as the difference between the two Lorenz curves:

$$\mathcal{M}(u) = \mathcal{L}^*(u) - \mathcal{L}_R^*(u) \quad (4.8)$$

Aaberge and Mogstad (2014) show that $\mathcal{M}(u) \geq 0$ for all $u \in [0; 1]$, *i.e.* that the Lorenz curve of the reference distribution is below the Lorenz curve of actual lifetime income (which means that it is more unequal).

Mobility curves can be synthesized into income mobility indices in the same way Lorenz curves can be synthesized into income inequality indices. For instance, an income mobility index can be obtained by integrating a weighting function along the income mobility curve:

$$\Lambda_p(\mathcal{M}) = \int_0^1 p(u) d\mathcal{M}(u) = \int_0^1 p(u) d(\mathcal{L}^* - \mathcal{L}_R^*)(u) = J_p(\mathcal{L}_R^*) - J_p(\mathcal{L}^*) \quad (4.9)$$

If we consider the same weighting function as for the Gini index ($p(u) = 2(1 - u)$), $\Lambda_p(\mathcal{M})$ is simply the difference between the Gini indices of both distribution. Aaberge and Mogstad (2014) argue that taking the difference rather than the ratio is preferable in that a ratio whose denominator is very small will be very sensitive to small changes in the numerator. It will be equal to zero if the actual distribution of lifetime income is the same as the reference distribution that corresponds to the perfect rigidity scenario. It will be equal to one if the actual lifetime income distribution is perfectly equal *and* if the reference distribution is completely unequal, *i.e.* if all income belongs to one individual only.

4.2.3 Measuring segregation

Segregation is the fact that people with different characteristics (*e.g.* race, gender, socioeconomic status) are isolated from each other. Its measurement raises a number of conceptual and technical issues (Massey and Denton, 1988). In our case, we choose to focus on segregation indices that measure the extent to which the population is evenly distributed across units (*e.g.* cities or schools) depending on their characteristics. This definition is associated with a number of popular segregation indices such as the dissimilarity index, the Gini index, the normalized

exposure and the entropy index. Frankel and Volij (2011) provide an excellent review of the main segregation indices, including their definitions and axiomatic properties.

We can see segregation indices as measures on an undirected graph where nodes are individuals and edges are social relations. We assume that each individual belongs to one of two groups denoted A and B.⁴ Group A will be called the *reference group* (e.g. the rich, the black, the women, etc.) and group B will contain everyone else in the population. For instance, if group A denotes black individuals, group B will contain all non-black individuals (not only whites). The reference group may be the majority group or a minority, indifferently: we call $p \in [0; 1]$ its share in the population. Each individual $i \in \{1, \dots, N\}$ belongs either to group A ($U_i = 1$) or group B ($U_i = 0$). The *social environment* of individual i is the set of individuals that are connected to i through an edge, *including individual i herself*.⁵ We denote μ_i the share of i 's social environment that belongs to the reference group: it is called the *exposure to the reference group* of individual i .

In the two-group case, segregation indices are essentially measures of the inequality of the distribution of μ_i , *i.e.* they are indices of social environment inequality. However, there is one important difference between income inequality and social environment inequality, which is that individuals *are* each other's social environment. Therefore, the distribution of μ_i is subject to constraints. In particular, a widespread assumption is that the graph consists of a set of isolated, complete subgraphs, *i.e.* that individuals interact in clusters, as shown in FIGURE 4.2. With this hypothesis, the clusters are denoted $k \in \{1, \dots, K\}$, each cluster contains N_k individuals and has a share p_k of individuals from the reference group. Therefore, if individual i belongs to cluster k , $\mu_i = p_k$. The average of μ_i over the whole population is then equal to p :

$$\mathbb{E}(\mu_i) = \frac{1}{N} \sum_{i=1}^N \mu_i = \frac{1}{N} \sum_{k=1}^K N_k p_k = p$$

This hypothesis is very common in the literature on segregation, and most segregation indices

4. Segregation indices can be defined with more than two groups, as detailed in Frankel and Volij (2011). We focus on the two-group case for simplicity's sake; adding more groups has no fundamental impact on the analysis.

5. The choice of including individual i in her own social environment or not is quite arbitrary: both options have advantages and drawbacks and they have their own justifications. The numerical difference is small when individuals have a large number of social relations. We choose to include individual i in her own social environment in order to simplify calculations.

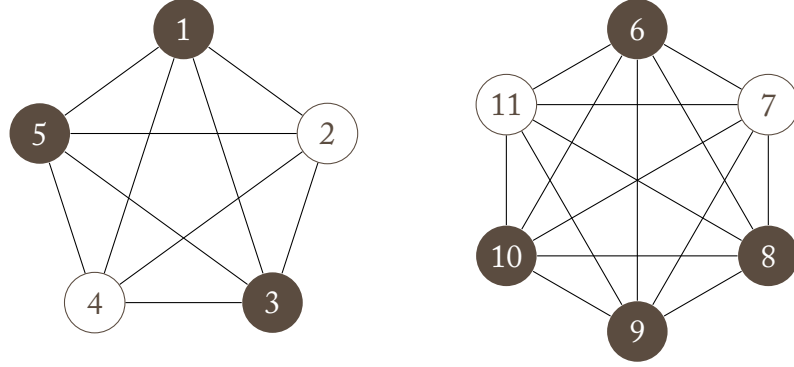


FIGURE 4.2 – A social network graph with interactions in clusters. Filled circles denote individuals from group A and plain circles denotes individuals from group B. The average exposure to the reference group is 60 percent in the left cluster, 67 percent in the right cluster.

are defined based on these clusters. Below are the definitions of four popular segregation indices: the dissimilarity index (D), the Gini index (G), the normalized exposure (P), and the entropy index (H).

$$D = \frac{1}{2p(1-p)} \sum_{k=1}^K \frac{N_k}{N} |p_k - p| \quad (4.10)$$

$$G = \frac{1}{2p(1-p)} \sum_{k=1}^K \sum_{k'=1}^K \frac{N_k}{N} \frac{N_{k'}}{N} |p_k - p_{k'}| \quad (4.11)$$

$$P = \frac{1}{p(1-p)} \sum_{k=1}^K \frac{N_k}{N} (p_k - p)^2 \quad (4.12)$$

$$H = 1 - \frac{1}{h_2(p)} \sum_{k=1}^K \frac{N_k}{N} h_2(p_k) \quad (4.13)$$

where h_2 is the 2-category entropy function, defined by:

$$h_2(u) = u \log_2 \left(\frac{1}{u} \right) + (1-u) \log_2 \left(\frac{1}{1-u} \right) \quad (4.14)$$

Note that the definition of the Gini index is similar to the Gini index on income distribution defined in equation 4.3.

All these indices equal zero when $p_k \equiv p$ in all clusters, *i.e.* if every cluster has the same share of individuals from the reference group: in that case, there is no segregation at all. Conversely,

they all equal one when $p_k \in \{0, 1\}$ in all clusters, *i.e.* if every cluster contains either only individuals from the reference group, or no individual from that group: there is no mixity in any of the clusters, segregation is complete.

The indices take different values for intermediate situations, they have different interpretations and different axiomatic properties, as detailed in [Frankel and Volij \(2011\)](#). For instance, the dissimilarity index D is equal to the minimal share of the population that would have to move to a different cluster in order to achieve full desegregation, normalized by its maximum value; the normalized exposure index P has three possible interpretations:

- First, it can be seen as the variance of p_k across all clusters, normalized by its highest possible value $p(1 - p)$ which is reached if each cluster contains either only individuals from group A or only individuals from group B.
- Second, P is the share of between-clusters variance of the binary variable U_i (belong to the reference group), since $p_k = \mathbb{E}(U_i | i \in k)$. It is the R^2 of the regression of that variable on a cluster-level fixed effect.
- Third, calling μ^A the average exposure to group A *of the individuals in that group* and μ^B the average exposure to group A *of the individuals from group B*, the value of the index is equal to:

$$P = 1 - \frac{\mu^B}{p} = \mu^A - \mu^B \quad (4.15)$$

i.e. P is the exposure gap to the reference group between the two groups (proof in [appendix 4.A.3](#)).

Note that since $P \geq 0$, these equalities yield the following double inequality:

$$\mu^B \leq p \leq \mu^A \quad (4.16)$$

The reference group always has a greater exposure to itself than average, while the rest of the population has a lower exposure to the reference group.

In [FIGURE 4.3](#), we consider the case of 8 students ($N = 8$) attending one of two classes ($K = 2$). Half of them are in the reference group A (full circles) and the other half are in the group B (plain circles): $p = 0.5$. The students in class X have an exposure to the reference group $p_X = 0.75$ while the students in class Y have an exposure $p_Y = 0.25$. The segregation indices equal $D = 0.5$, $G = 0.5$, $P = 0.25$ and $H = (3/4) \log_2(3) - 1 \simeq 0.19$.

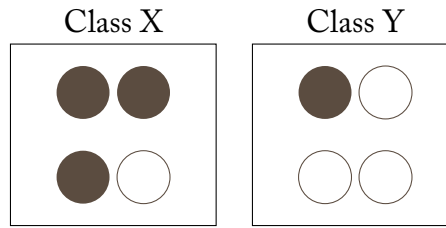


FIGURE 4.3 – Segregation between upper class (full circles) and middle or lower class (plain circles) students.

4.3 Social environment mobility

The tools presented in the previous section lay out the grounds for building social environment mobility (SEM) indices. As seen in section 4.2.3, segregation indices are not straightforward adaptations of income inequality indices. Similarly, income mobility indices cannot be transposed directly to the case of SEM. In this section, we detail the fundamental differences and the adaptations they imply on the definition of SEM indices. For the sake of simplicity, we take the example of social segregation between upper class individuals (group A) and middle or lower class individuals (group B). We sometimes refer to the group B as lower class individuals, omitting the fact that group B contains both middle class and lower class individuals.

4.3.1 Breaking out of the clusters

While the definitions of the segregation indices defined in section 4.2.3 are convenient, they are not appropriate when studying SEM. Their main drawback is that they rely on the fact that individuals interact in clusters, which makes sense when looking at segregation at one point in time, but not when looking at SEM. For instance, in the case of school segregation, clusters may represent classes in a school, and those are not constant across school years. Individuals interact in clusters in each school year, but not overall: if student A interacted with B and B interacted with C, A may not have interacted with C.

In FIGURE 4.4, students 1 to 8 are assigned to classes X and Y in the first year, and classes X' and Y' in the second year. In case 1, classes X' and Y' are identical to classes X and Y, but they are not in case 2. If we want to provide some information on the variety of social environments

students have experienced, the clusters only define the students' social environments *at one point in time*. But belonging to different clusters over time is an indication of social environment mobility. Just like income mobility indices, SEM indices provide more additional information on the degree of segregation over time. Although segregation indices have the same values in case 1 and in case 2, case 2 leads to less segregation over time: for instance, student 4 was the only lower class student in her class two years in a row in case 1 but not in case 2.

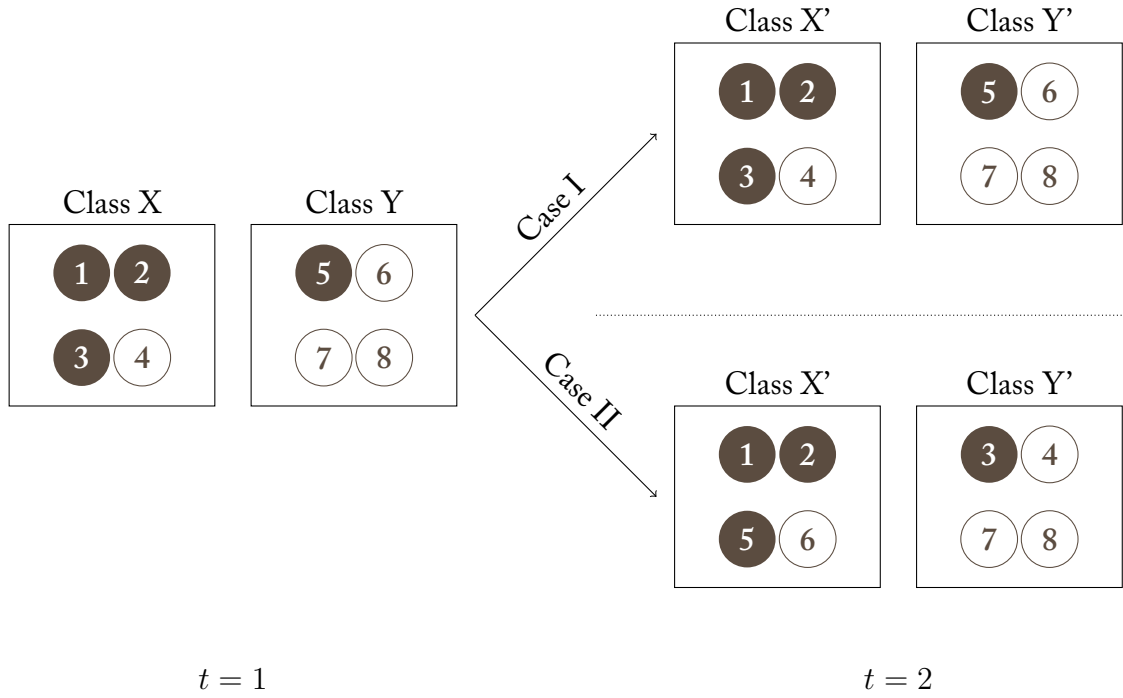


FIGURE 4.4 – Social environment mobility in the case of school segregation: students 1, 2, 3 and 5 are the group A (upper class students) and students 4, 6, 7 and 8 are the group B (lower class students). In the first case, the two classes are identical in both time periods while they are reshuffled in the second case, leading to a higher SEM.

Therefore, we need to define measures of segregation that do not rely on clusters. Fortunately, the indices introduced in the previous section can be written as a function of the individual exposures to the reference group μ_i . As we introduce a time variable $t \in \{1, \dots, T\}$, we will now denote μ_{it} the social environment of individual i at time t (which is still computed using the cluster individual i belongs to at time t). For now, we assume that the population is stable over time and that the individual characteristics that define the reference group are stable as well:

therefore, N and p are constants.

$$\begin{aligned}
 D_t &= \frac{1}{2p(1-p)} \frac{1}{N} \sum_{i=1}^N |\mu_{it} - p| \\
 G_t &= \frac{1}{p(1-p)} \frac{1}{N^2} \sum_{i=1}^N \sum_{i'=1}^N |\mu_{it} - \mu_{i't}| \\
 P_t &= \frac{1}{p(1-p)} \frac{1}{N} \sum_{i=1}^N (\mu_{it} - p)^2 \\
 H_t &= \frac{1}{h(p)} \frac{1}{N} \sum_{i=1}^N h(\mu_{it})
 \end{aligned}$$

With these expressions, segregation indices are a measure of inequality on the distribution of the social environments μ_{it} . In section 4.2.1, we introduced income inequality indices that were functions of the Lorenz curve. Yet the literature on segregation has not been using the Lorenz curve as such, but rather an adaptation of it introduced by [Duncan and Duncan \(1955\)](#) and called the *segregation curve*. However, the definition of this curve requires that individuals interact in clusters. The segregation curves is a continuous, piecewise linear function (*i.e.* a spline function of degree one) which goes through the points $(X_k, Y_k)_{k \in \{0, \dots, K\}}$ where $(X_0, Y_0) = (0, 0)$; X_k is the share of the group A that belongs to the k first clusters with the smallest share of individuals from that group; Y_k is the share of the group B that belongs to those same k clusters. It follows that $(X_K, Y_K) = (1, 1)$. The segregation curve has interesting properties, in particular it has geometric relations to some segregation indices, notably the dissimilarity index and the Gini index. The dissimilarity index is the maximal vertical distance between the curve and the 45-degree line (whose highest possible value is one). The Gini index is the area between the curve and the 45-degree line, divided by its highest possible value $1/2$: this is the same relation as the Gini index and the Lorenz curve in the case of income inequality.

Since we cannot consider clusters when studying SEM, we choose to go back to the initial Lorenz curve. To that end, we introduce the cumulative distribution function of μ_{it} , denoted

$\mathcal{F}_t$. Our four segregation indices can be written in integral form using $\mathcal{F}_t$:

$$D_t = \frac{1}{2p(1-p)} \int_0^1 |\mathcal{F}_t^{-1}(u) - p| \, du \quad (4.17)$$

$$G_t = \frac{1}{2p(1-p)} \int_0^1 \int_0^1 |\mathcal{F}_t^{-1}(u) - \mathcal{F}_t^{-1}(u')| \, du' \, du \quad (4.18)$$

$$P_t = \frac{1}{p(1-p)} \int_0^1 (\mathcal{F}_t^{-1}(u) - p)^2 \, du \quad (4.19)$$

$$H_t = \frac{1}{h(p)} \int_0^1 h(\mathcal{F}_t^{-1}(u)) \, du \quad (4.20)$$

The Lorenz curve for social environments at time t , denoted $\mathcal{L}_t$ has the same definition as the Lorenz curve for the income distribution. When there is no segregation, the Lorenz curve coincides with the 45-degree line. However, when segregation is maximal, $\mathcal{L}_t(u) = 0$ only if $u \leq 1 - p$. Recall that when all the income is earned by only one individual, $\mathcal{L}_t(u) = 0$ if $u < 1$. In the case of maximal segregation, the group A (which represents a share p of the population) is exposed to itself and to itself only, therefore they have $\mu_{it} = 1$, while the group B (whose share is $1 - p$) has $\mu_{it} = 0$. As a consequence, the Lorenz curve for social environments has a lower envelope which is equal to zero on $[0, 1 - p]$ and is linear on $[1 - p, 1]$, with $\mathcal{L}_t(1) = 1$, as shown on [FIGURE 4.5](#) (blue curve).

With this limitation, the highest possible vertical distance between the Lorenz curve and the 45-degree line is $1 - p$ and the highest possible area is $(1 - p)/2$. Interestingly, the dissimilarity index defined above is equal to the maximum vertical distance between the Lorenz curve and the 45-degree line divided by the highest possible value $1 - p$. Similarly, the Gini index is equal to the area between the Lorenz curve and the 45-degree line divided by the highest possible value $(1 - p)/2$. Therefore, the Lorenz curve remains useful when studying segregation as long as we keep in mind that it has a lower envelope that affects the normalization of the indices.

In order to adapt income mobility measures that were introduced in [section 4.2.2](#), we also need to introduce the distributions of *lifetime exposure*, i.e. of the variable $\bar{\mu}_i = (1/T) \sum_{t=1}^T \mu_{it}$. We call $\mathcal{F}^*$ the corresponding cumulative distribution function and $\mathcal{L}^*$ the Lorenz curve. We can then define the *lifetime segregation indices* S^* (where S is any of the four segregation indices introduced in [section 4.2.3](#)) by replacing the instantaneous distribution $\mathcal{F}_t$ with the lifetime

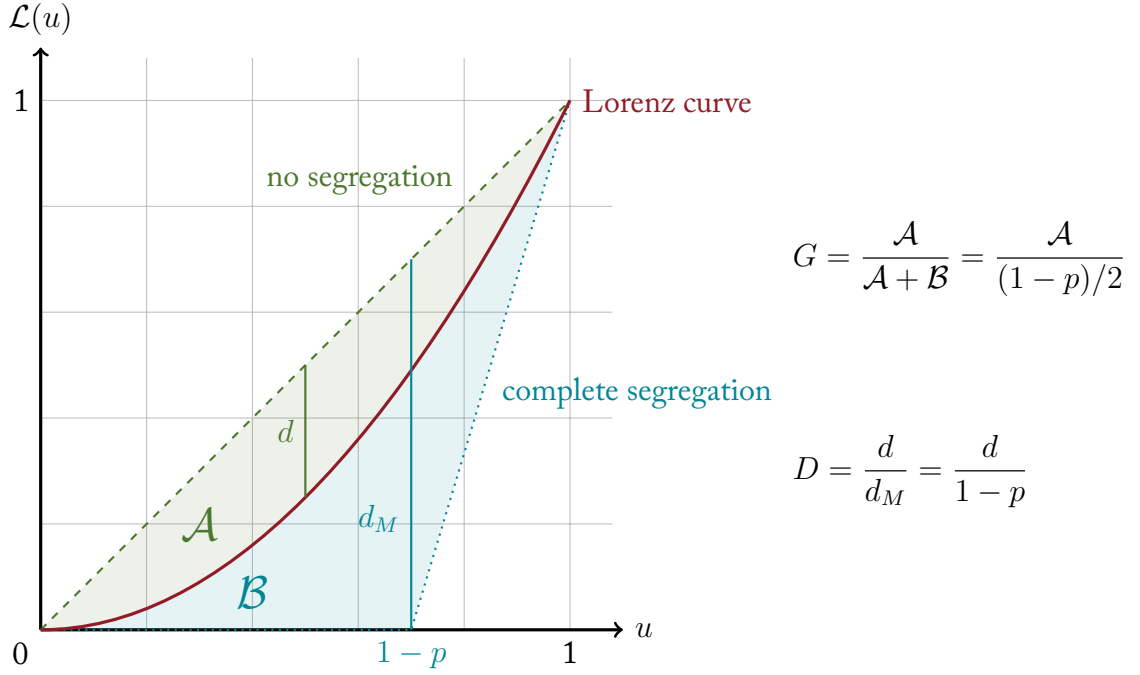


FIGURE 4.5 – The Lorenz curve for social environments

distribution $\mathcal{F}^*$ in the formulae (4.17), (4.18), (4.19) or (4.20). Note that unlike instantaneous segregation indices, lifetime segregation indices generally do not have simple interpretations, except for the lifetime normalized exposure $\bar{P}$ which is the gap in average exposure to the reference group between the two groups A and B: $\bar{P} = \bar{\mu}^A - \bar{\mu}^B$.

We then have all the tools to adapt the income mobility measures to the case of SEM:

- Following [Shorrocks \(1978a\)](#), we can define SEM indices that compare the lifetime segregation index with the average instantaneous segregation index:

$$\mathcal{M}^S = 1 - \frac{S^*}{\bar{S}} \quad (4.21)$$

where $\bar{S}$ is the average instantaneous segregation, *i.e.* the average value of S_t over the T periods. Note that for any segregation index, $S^* \leq \bar{S}$, so that $\mathcal{M}^S$ belongs to the unit interval. This is true for the same reasons why the income mobility index defined in section 4.2.2.2 belongs to the unit interval, since the mathematical definition is identical (see [Shorrocks, 1978a](#), p. 386).

- [Aaberge and Mogstad \(2014\)](#)'s income mobility curves can also be imported to social

environment mobility using the exact same definition:

$$\mathcal{M}(u) = \mathcal{L}^*(u) - \mathcal{L}_R^*(u) \quad (4.22)$$

where the Lorenz curve $\mathcal{L}_R^*$ of the "reference distribution" that describes the perfect rigidity scenario has the same definition as in the income mobility case. Both the lifetime segregation Lorenz curve and the reference curve have the same lower envelope, therefore the difference has no lower envelope other than the zero function.

Although the definitions are identical, the SEM indices have one important difference with income mobility indices, which is that even when SEM is very high, lifetime segregation can never vanish to zero as long as there is some degree of instantaneous segregation. We examine this point in more details in the next section.

4.3.2 The limited equalizing force of social environment mobility

As seen in section 4.2.2, income mobility gives a different perspective on income inequality, as a significant fraction of income inequality may vanish when comparing lifetime incomes instead of instantaneous incomes. In fact, theoretically, there could exist a society in which instantaneous inequality is always high but income mobility may be so high that inequality in lifetime incomes equals zero. This would be the case, for instance, if the income was a direct function of the individual's age and all individuals have the same lifespan.

Yet in the case of SEM, it is not possible to have no lifetime segregation if we have instantaneous one. As stated in equation (4.16), the group A will always have a greater exposure to itself than the group B ($\mu_t^A \geq \mu_t^B$), unless there is no segregation in which case the exposures are equal. Suppose that in a given period, there is some segregation: then, for that period, group A will have a greater exposure to itself than group B. If we want lifetime segregation to be zero, *i.e.* if we want all individuals to have the same lifetime exposure to the group A, then in average the group B must have a greater exposure to the group A than the group A itself in other periods: this cannot happen. Therefore, any level of instantaneous segregation has irreversible consequences on the minimum level of lifetime segregation that may be observed after several

time periods. An implication is that the Lorenz curve of lifetime exposures $\bar{\mathcal{L}}(u)$ cannot coincide with the 45-degree line: instead, it has an upper envelope $\bar{\mathcal{L}}^*(u)$. Our goal in this section is to determine this upper bound, *considering instantaneous segregation as given*. In other words, we are interested in finding the upper envelope $\bar{\mathcal{L}}^*(u)$ for a given series of the instantaneous distributions $(\mathcal{F}_t)$.

The intuition is the following. As explained in section 4.2.3, as long as there is some level of segregation at any time, white people and non-white people cannot have the same exposure to white people. The values of μ_t^A and μ_t^B are determined by the level of instantaneous segregation, which we consider as given. More precisely, using the normalized exposure index P_t , we have the relationships:

$$P_t = 1 - \frac{\mu_t^B}{p} = \mu_t^A - \mu_t^B \quad (4.23)$$

Since these relationships are linear, we can average over time:

$$\bar{P} = 1 - \frac{\bar{\mu}^B}{p} = \bar{\mu}^A - \bar{\mu}^B \quad (4.24)$$

Now, in the best case scenario, there is no difference in $\bar{\mu}_i$ between individuals from the reference group and between individuals from the group B, i.e. $\bar{\mu}_i = \bar{\mu}^g$ if $i \in g$ where g denotes group A or B: if SEM is optimal, it will eliminate differences within each group, but it cannot eliminate differences between the two groups.

In that case – this is our main result – the Lorenz curve of lifetime social environments is a spline function of degree one with a knot at $u = 1 - p$ with $\mathcal{L}^*(1 - p) = (1 - p)(1 - \bar{P})$, as plotted in [FIGURE 4.6](#) (see proof in appendix 4.A.4).

This upper envelope translates into lower bounds for the four lifetime segregation indices:

$$D_m^* = \bar{P} \quad (4.25)$$

$$G_m^* = \bar{P} \quad (4.26)$$

$$P_m^* = \bar{P}^2 \quad (4.27)$$

$$H_m^* = \frac{(1 - p) \cdot h(p(1 - \bar{P})) + p \cdot h(p(1 - \bar{P}) + \bar{P})}{h(p)} \quad (4.28)$$

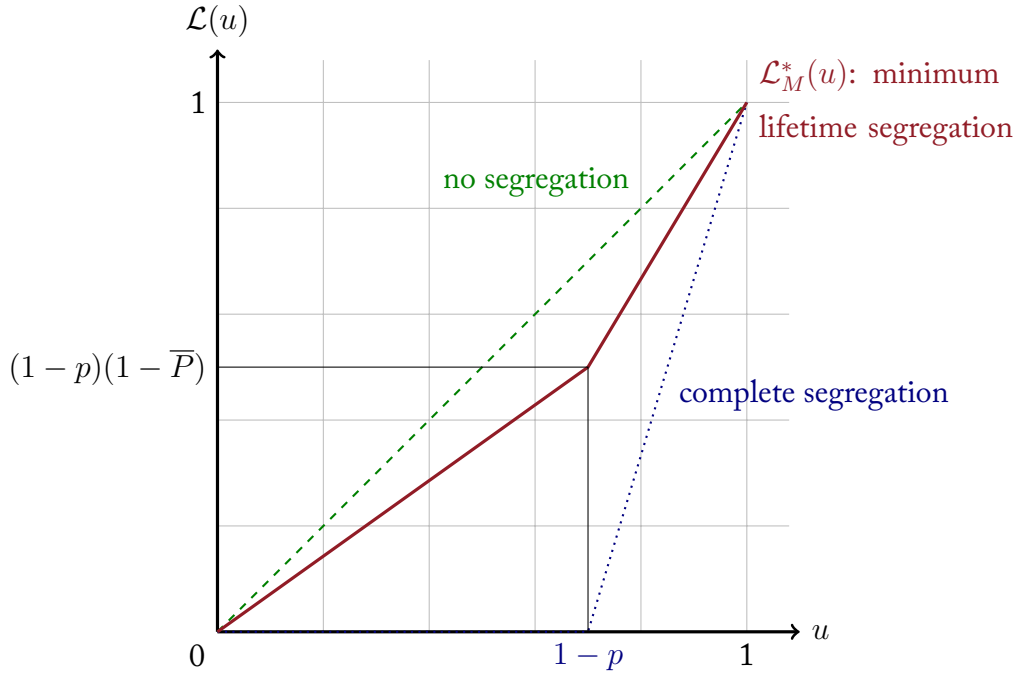


FIGURE 4.6 – The upper envelope of lifetime social environment Lorenz curves

and into upper bounds for the SEM indices adapted from [Shorrocks \(1978a\)](#):

$$\mathcal{M}_M^D = 1 - \frac{\bar{P}}{\bar{D}} \quad (4.29)$$

$$\mathcal{M}_M^G = 1 - \frac{\bar{P}}{\bar{G}} \quad (4.30)$$

$$\mathcal{M}_M^P = 1 - \bar{P} \quad (4.31)$$

$$\mathcal{M}_M^H = 1 - \frac{H_m^*}{\bar{H}} \quad (4.32)$$

Note that this upper bound to SEM indices is a theoretical one. Reaching this upper bound exactly is not always possible if the number of time periods is not big enough. The SEM index is equal to its upper bound if and only if all individuals within each group have the exact same lifetime exposure to the reference group.

Consider the case described in [FIGURE 4.3](#) and suppose that each year, there are two classes with identical group compositions but possibly different individual compositions (*i.e.* class X always contains three students from group A but they can be any of the four group A students).

The value of the segregation index P_t will be 25 percent in each year, therefore $\bar{P}$ is also equal to 25 percent and the upper bound of the SEM index is 75 percent as per equation (4.31). The average exposure of the upper class students is equal to 62.5 percent in each period: if SEM is complete, the lifetime exposure of each upper class student should be equal to 62.5 percent. This is possible if and only if each student spends three time periods in class X for each time period spent in class Y: therefore, the number of periods must be a multiple of four (there is an identical constraint on lower class students).

Therefore, the number of time periods induces a bias on the upper bound to mobility. This bias decreases with the number of time periods: in our example, complete mobility could not be reached for $T = 10$ or $T = 14$ but the variance of lifetime exposure within *e.g.* the reference group would be smaller in the second case. The bias also gets smaller when the distribution of exposures at each time is peaked around its mean value: in that case, a large share of individuals will be close to the mean value at each time and the lifetime exposure can converge more rapidly towards this mean value.

Unfortunately, there is no simple formula for this bias: finding the optimal series of allocations of individuals across clusters at each time is a complicated algorithm. In practice however, if the allocation of individuals across clusters is not too peculiar, the distribution of exposures allows for a relatively rapid convergence towards a zero bias.

4.3.3 Dealing with attrition

Attrition is a classic issue when working with panel data. In the case of SEM, attrition entails an additional methodological difficulty. While we need individuals to be observed several times in order to compute a mobility index, individuals who are observed only once still affect the other individuals' social environments. Therefore, it is not possible to simply remove individuals who are not always observed out of the sample: at least, they must be taken into account when computing the exposures.

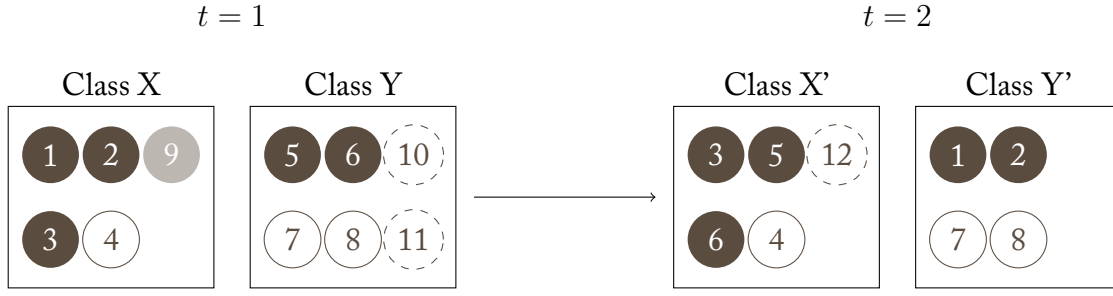


FIGURE 4.7 – Attrition: students 9, 10, 11 and 12 are not present in both time periods but they cannot be ignored since they affect the social environment of students 1 to 8.

Call Ω_t the set of individuals observed in t and define:

$$\Omega = \bigcup_{t=1}^T \Omega_t \quad \text{and} \quad \omega = \bigcap_{t=1}^T \Omega_t$$

Ω is the set of individuals that we observe at least once and $\omega \subset \Omega$ is the *balanced panel*, *i.e.* the subset of individuals that we observe in each time period. The larger sample Ω is used to compute the exposures to the reference group: the social environment of individual i at time t is a subset of Ω (but maybe not of ω), and μ_{it} is the share of individuals from the reference group in that subset. For instance, **FIGURE 4.7** represents a group of 12 students who interact in two classes at two points in time. The restricted sample is $\omega = \{1, 2, 3, 4, 5, 6, 7, 8\}$ and the full sample is $\Omega = \omega + \{9, 10, 11, 12\}$. In the first period, the social environment of student A is the entire class X, which contains 5 students: 4 upper class students and one lower class student. Therefore, the student 1 has exposure $\mu_{1,1} = 4/5 = 0.8$.

We could then compute instantaneous segregation indices based on the distribution of μ_{it} on the restricted sample, using the formulae (4.17), (4.18), (4.19) and (4.20). However, these definitions may lead to segregation indices out of the unit interval. This also means that the SEM indices based on segregation indices cannot be used directly.

In this section, we focus on the SEM index $\mathcal{M}^P$, which is based on the normalized exposure segregation index and which can be easily adapted to a case with attrition – unlike the other SEM indices that cannot be adapted in a simple way.

As stated earlier, the normalized exposure index P_t is equal to the variance of the series of exposures $(\mu_{it})_{it \in \{1, \dots, N\} \times \{1, \dots, T\}}$, normalized by its highest possible value $p(1 - p)$, which would be reached if each class contained either only individuals from group A or only individuals from group B. Similarly, P^* is the variance of the series $(\bar{\mu}_i)_{i \in \{1, \dots, N\}}$, normalized by its highest possible value which is also $p(1 - p)$. Since the normalization is the same for P_t and P^* , the ratio $P^*/\bar{P}$ is also equal to the ratio of the variances $V^* = p(1 - p)P^*$ and $V = p(1 - p)\bar{P}$. V^* is the variance of the lifetime exposures $\bar{\mu}_i$ while V is the total variance of μ_{it} (as long as p is constant, which is one of our hypotheses). Then,

$$\mathcal{M}^P = 1 - \frac{V^*}{V} = \frac{V - V^*}{V} \quad (4.33)$$

This new definition of the mobility index $\mathcal{M}^P$ has an interesting interpretation. Indeed, a variance decomposition shows that $V - V^*$ is equal to the average individual variance of exposure across time periods: if v_i is the variance of the series $(\mu_{it})_t$, $V - V^*$ is the average value of v_i . Therefore, $V - V^*$ captures the diversity of the social environments experienced by individuals. The SEM index $\mathcal{M}^P$ compares the average diversity of social environments experienced by a single individual to the diversity of all existing social environments in the setting. An equivalent approach is to look at the panel series (μ_{it}) and to regress it against an individual fixed effect:

$$\mu_{it} = \alpha_i + \epsilon_{it} \quad (4.34)$$

$\mathcal{M}^P$ is then equal to $1 - R^2$ where R^2 is the coefficient of determination of the above regression.

The definition can then be adapted to a scenario with attrition. Denoting V_ω^* the variance of the series $(\bar{\mu}_i)_{i \in \omega}$ and V_ω the variance of the series $(\mu_{it})_{it \in \omega \times \{1, \dots, T\}}$, we can use this simple adaptation of the SEM index:

$$\mathcal{M}_\omega^P = 1 - \frac{V_\omega^*}{V_\omega} = \frac{V_\omega - V_\omega^*}{V_\omega} \quad (4.35)$$

i.e. we compare the average individual variance to the overall variance of the exposures *in the reference group*. The interpretation given above is still valid and $\mathcal{M}^P$ still belongs to the unit interval.

The formula of the upper bound defined in section 4.3.2 should also be adapted to this new definition. In order to compute this upper bound, we consider once again the scenario of optimal

SEM. In this best case scenario, all individuals from the reference group *in the balanced panel* have the same lifetime exposure, and all individuals from the group B *in the balanced panel* has the same lifetime exposure too. These two values $\overline{\mu_\omega^A}$ and $\overline{\mu_\omega^B}$ can be computed directly from the actual distribution of μ_{it} :

$$\overline{\mu_\omega^g} = \mathbb{E}(\mu_{it} | i \in \omega \cap g) \quad (4.36)$$

Denoting p_ω the share of the reference group in the balanced panel and $\overline{\mu_\omega}$ the average exposure to the reference group in this sample,⁶ the smallest possible variance of lifetime exposures is given by

$$V_{\omega,m}^* = p_\omega(\overline{\mu_\omega^A} - \overline{\mu_\omega})^2 + (1 - p_\omega)(\overline{\mu_\omega^B} - \overline{\mu_\omega})^2 \quad (4.37)$$

Therefore, the upper bound of the SEM index is:

$$\mathcal{M}_{\omega,M}^P = 1 - \frac{p_\omega(\overline{\mu_\omega^A} - \overline{\mu_\omega})^2 + (1 - p_\omega)(\overline{\mu_\omega^B} - \overline{\mu_\omega})^2}{V_\omega} \quad (4.38)$$

4.3.4 Nested levels of clustering

When studying segregation, it is not uncommon to find cases where there are nested levels of clustering. The typical example is school segregation, where individuals (students) interact mainly within classes but also within schools that contain several classes. The value of the segregation index varies depending on the choice of the level of clustering, *i.e.* on the definition of a student's social environment. In particular, if we take the classroom as the reference cluster, the segregation index will be larger than if the reference cluster is the whole school. In the case of the normalized exposure index P , there is a simple relationship between the indices P^{class} and P^{school} .

Call μ_i^{class} (resp. μ_i^{school}) the share of the reference group in student i 's class (resp. school). Recall that U_i is a binary variable that equals 1 if the individual i belongs to the reference group. As stated earlier, the segregation index P^{class} is the share of the variance of U_i that is explained by classes, which is equal to the variance V^{class} of the series $(\mu_i^{\text{class}})_i$ divided by the variance of U_i , *i.e.* by $p(1 - p)$. Similarly, the segregation index P^{school} is equal to the variance V^{school} divided

6. $\overline{\mu_\omega}$ is different from p_ω because the individuals from the balanced panel interact with individuals from Ω who are not in ω . If there was no attrition ($\omega = \Omega$), then we would have $\overline{\mu_\omega} = p_\omega$.

by $p(1 - p)$. Note that the series $(\mu_i^{\text{class}})_i$ (resp. $(\mu_i^{\text{school}})_i$) is identical to the series of the shares of the reference group in each class (resp. school), weighting each class (resp. school) by its size.

Therefore, the total variance $p(1 - p)$ is the sum of three terms:

1. the variance of μ_i^{school} , which represents the *between-school variance*;
2. the average *within-school variance* of μ_i^{class} ;
3. the average within-class of the variable U_i , which is the remaining variation that is not explained by schools or classes.

The share of the first term in the total variance is the between-school segregation P^{school} . The share of the first two terms is the between-class segregation P^{class} . This reasoning can be extended to more clustering levels: city, region, etc.

Multi-level clustering is also interesting in the case of SEM. In section 4.3.1, we explained how the segregation indices had to be reworked in order to be independent of the notion of clusters, since SEM implies that individuals do not remain in the same cluster throughout their life. However, when there are several levels of clustering, we can imagine that the smallest-level clusters are indeed unstable over time, but that larger clusters do remain throughout all time periods. In the case of school segregation, we could imagine that classes are reshuffled every year but that students cannot change schools. In this scenario, the SEM index has to be adapted.

As detailed in section 4.3.3, the SEM index $\mathcal{M}^P$ can be seen as the ratio of the average individual variance of exposure to the variance of all exposures in the population and throughout all time periods. Yet if students cannot leave their school, it makes sense to compare the diversity of exposures experienced by an individual student to the distribution of exposures that are available in their school rather than in any existing school. There are two ways in which this can be taken into account in the SEM indices:

1. If students cannot change schools, the upper bound to mobility is actually lower than the standard value. The best case scenario under this constraint is that *within each school, every individual from the reference group has the same lifetime exposure and every individual from*

the other group has the same lifetime exposure. The value of the upper bound of the SEM index can easily be computed under this condition.

2. It may also make sense to change the denominator of the SEM index given in equation (4.33) in order to compare the individual variances to the within-school variance rather than the total variance V . Note that in this case, the upper bound should also be computed using the method described in the previous point (optimal within-school mobility).

In practice however, clusters are rarely perfectly persistent over time: although most students remain in the same school, a small fraction of them is likely to change schools. The existence of such students has important consequences on the within-school SEM index's axiomatic properties: in theory, if some students may change schools, within-school mobility may exceed one. If these students are a small minority, we choose to treat them as if they left the balanced panel and we consider them as attrition: in the notations of section 4.3.3, they belong to $\Omega \setminus \omega$, *i.e.* they do affect the other students' environments, but they are not taken into account when computing the SEM index.

Overall, if we consider two levels of clustering, there are three SEM indices to consider:

1. the standard SEM index where we compare the average individual variance to the overall variance in exposures;
2. the same standard SEM index but restricted to a balanced panel consisting of individuals who remain in the same school;
3. the SEM index obtained by dividing the average individual variance of the restricted balanced panel by the overall within-school variance.

The difference between the values of the first and the second index should be small and is only due to a selection effect. However, they have different upper bounds, since in the second case the relevant upper bound is based on the assumption that students cannot change schools. The third index will be mechanically higher than the second, since the only difference is to replace the total variance by the within-school variance in the denominator, in both the SEM index and the upper bound.

4.4 Case study: SEM in French middle schools

In this section, we apply the SEM analysis methodology that we laid out in section 4.3 to a concrete example. All segregation and SEM indices rely on the normalized exposure index P , which has several advantages:

- the lifetime segregation index P^* has an intuitive interpretation, unlike other lifetime segregation indices;
- the upper bound to the SEM index can be expressed as a function of the average instantaneous segregation $\bar{P}$ when there is no attrition;
- when there is attrition, we can use adapted versions of the SEM indices based on $\mathcal{M}^P$;
- we can use the additive decomposability property of the segregation index and compare SEM indices on different levels.

We take the example of social segregation and SEM in French middle schools. In 2009, a cohort of 797,894 students entered 6th grade. These students spend four years in middle school, from grade 6 to grade 9. The objective of this section is to provide figures on social segregation in French middle schools and on the diversity of social environments experienced by the middle school students during these four years.

Thanks to a unique database from the French Ministry of Education, we are able to follow 79 percent of the cohort throughout the four years of middle school (possibly in different schools at different points in time): these students are the balanced panel. We know about their socioeconomic status and the class they attended in each year. The 21 percent of students who are not observed all four years either repeated a grade (7 percent) or could not be matched from one dataset to another (14 percent).

Our reference group is the upper class students, which make up 38 percent of the sample ($p = 0.38$); while the rest of the sample contains both lower class and middle class students, we refer to them as the lower class students for the sake of simplicity.

FIGURE 4.8 gives the value of the instantaneous segregation index at each grade $t \in \{6, 7, 8, 9\}$. In average, the segregation index equals 19 percent. Given that the share of up-

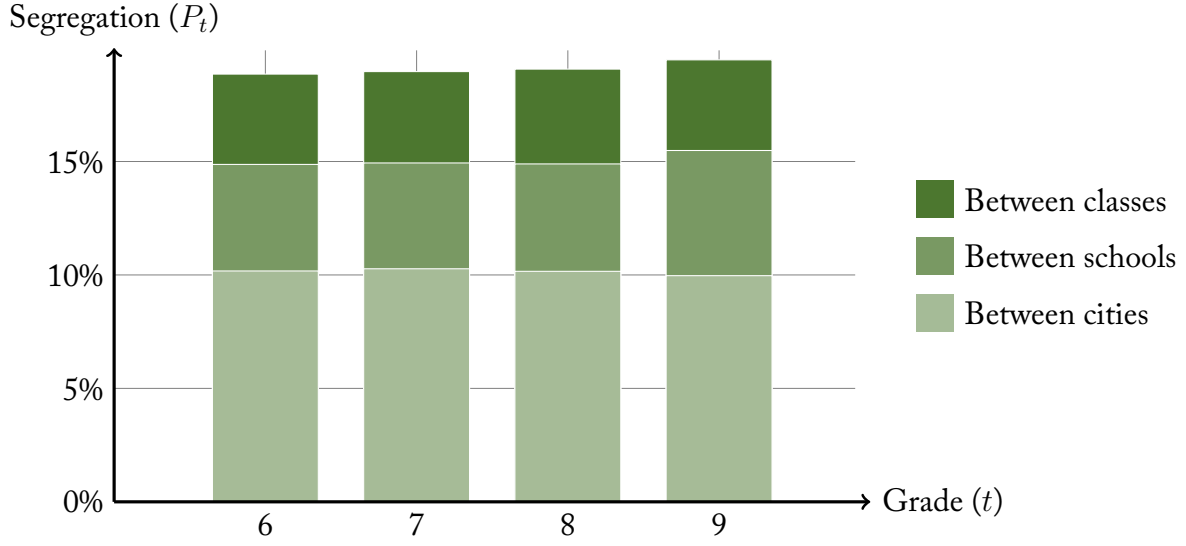


FIGURE 4.8 – Normalized exposure in French middle schools, cohort 2009

per class students is 38 percent, this means that in average, *the class of an upper class student contains 50 percent of upper class students, while the class of a lower class student contains only 31 percent of upper class students*, as per equation (4.15). The figures also show the decomposition of the segregation across three levels of clustering: cities, schools and classes.

The between-cities segregation is equal to 10 percent, which means that if all schools and all classes were identical within each city, the total segregation would equal 10 percent instead of 19 percent: residential segregation accounts for 53 percent of the total segregation. The other half is driven by within-city, between-school segregation (5 points, 26 percent of the total) and within-school, between-class segregation (4 points, 21 percent of the total).

Because of the attrition in our sample, we need to use the SEM index defined in equation (4.35) and the upper bound defined in equation (4.38). We find that the value of the SEM index $\mathcal{M}_\omega^P$ is 15 percent and its upper bound is 81 percent. The numerical value of the upper bound is very close to $1 - \bar{P}$ which is the value in the case where there is no attrition.

A simple, yet approximate, interpretation of the SEM index is the following. For each student in the balanced panel, compute the average share of upper class students in their class

throughout the four years of middle school. The gap between an upper class and a lower class student's lifetime exposure to upper class students, P^* , is about 15 percent less than the segregation index, *i.e.* 16 percent instead of 19 percent. This interpretation is not strictly correct since the numerator in equation (4.35) does not relate exactly to the exposure gap $\overline{\mu}_w^1 - \overline{\mu}_w^0$ (it does only if there is no attrition). In this case however, this interpretation remains fairly accurate – which can be checked by computing the actual lifetime exposures.

A second interpretation is that the individual variance of exposures to upper class students is about 15 percent of the variance of exposure throughout all classes in middle school. In other words, each student only experiences about 15 percent of the diversity of social environments available in all French middle schools. Although the average student could not experience all this diversity because of the existing segregation at each time, the actual value is quite far from the upper bound (81 percent).

Among the 630,466 from the balanced panel, 545,250 (86 percent) remained in the same school all four years. This indicates that mobility across schools is very limited. Therefore, it makes sense to measure the within-school SEM index. To that end, we focus on the restricted balanced panel of the 545,250 students who remained in the same school all four years. The SEM index is slightly lower for this sample at 12 percent instead of 15 percent. If SEM within each school was optimal, *i.e.* if all upper class students and all lower class students within each school had the same lifetime exposure, this SEM index would reach 20 percent: as long as students do not change schools, the social segregation is such that middle school principals could not get their students to experience more than 20 percent of the variety of the social environments in the country. If we compare the individual variances to the within-school variance instead of the total variance, we obtain the within-school SEM index, which equals 58 percent, with an upper bound of 93 percent. Overall, within-school mobility is quite substantial: this indicates that a large share of students experiences a fair share of the social environments that exist in the school throughout their four years.

Future work will provide a more in-depth analysis of segregation and social environment mobility, which this case study is a preview of. We also plan to analyze social environment

mobility in different contexts (*e.g.* residential mobility in France or in other countries).

4.5 Bibliography

- Aaberge, R. (2001). Axiomatic characterization of the gini coefficient and lorenz curve orderings. *Journal of Economic Theory*, 101(1):115–132.
- Aaberge, R. and Mogstad, M. (2014). Income mobility as an equalizer of permanent income. Discussion Paper 769, Statistics Norway Research department.
- Bonhomme, S. and Robin, J.-M. (2009). Assessing the equalizing force of mobility using short panels: France, 1990–2000. *The Review of Economic Studies*, 76(1):63–92.
- Bowlus, A. J. and Robin, J.-M. (2004). Twenty years of rising inequality in u.s. lifetime labour income values. *The Review of Economic Studies*, 71(3):709–742.
- Braddock II, J. H. and McPartland, J. M. (1989). Social-psychological processes that perpetuate racial segregation: The relationship between school and employment desegregation. *The Journal of Black Studies*, 19(3):267–289.
- Camargo, B., Stinebrickner, R., and Stinebrickner, T. (2010). Interracial friendships in college. *Journal of Labor Economics*, 28(4):861–892.
- Crain, R. L. and Strauss, J. (1985). School desegregation and black occupational attainments: Results from a long-term experiment. Technical Report 359, Center for the Social Organization of Schools, Baltimore.
- Currarini, S., Jackson, M. O., and Pin, P. (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica*, 77(4):1003–1045.
- Cutler, D. M. and Glaeser, E. L. (1997). Are ghettos good or bad? *The Quarterly Journal of Economics*, 112(3):827–872.
- Duncan, O. D. and Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, 20(2):210–217.

- Echenique, F. and Fryer, R. G. (2007). A measure of segregation based on social interactions. *The Quarterly Journal of Economics*, 122(2):441–485.
- Frankel, D. M. and Volij, O. (2011). Measuring school segregation. *Journal of Economic Theory*, 146(1):1–38.
- Gottschalk, P. and Spolaore, E. (2002). On the evaluation of economic mobility. *The Review of Economic Studies*, 69(1):191–208.
- Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, 9(70):209–219.
- Marmaros, D. and Sacerdote, B. (2006). How do friendships form? *The Quarterly Journal of Economics*, 121(1):79–119.
- Massey, D. S. and Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces*, 67(2):281–315.
- Moffitt, R. A. and Gottschalk, P. (2002). Trends in the transitory variance of earnings in the united states. *Economic Journal*, 112(478):C68–C73.
- Prais, S. J. (1955). Measuring social mobility. *Journal of the Royal Statistical Society. Series A (General)*, 118(1):56–66.
- Shorrocks, A. (1978a). Income inequality and income mobility. *Journal of Economic Theory*, 19(2):376–393.
- Shorrocks, A. (1978b). The measurement of mobility. *Econometrica*, 46(5):1013–1024.
- Wells, A. S. and Crain, R. L. (1994). Perpetuation theory and the long-term effects of school desegregation. *Review of Educational Research*, 64(4):531–555.
- Yaari, M. E. (1988). A controversial proposal concerning inequality measurement. *Journal of Economic Theory*, 44(2):381–397.

4.A Mathematical proofs

4.A.1 Equivalence of the two definitions of the Gini index

We prove here that the definitions of the Gini index provided in equations (4.2) and (4.3) are equivalent, *i.e.*:

$$G = 1 - 2 \int_0^1 \mathcal{L}(u) \, du = \frac{1}{2N^2} \sum_{i=1}^N \sum_{j=1}^N |y_i - y_j| \quad (4.39)$$

Without losing generality, we can assume that $i < j$ implies $y_i < y_j$, *i.e.* that the incomes are sorted in ascending order. Under that assumption, the double sum from the right hand side can be rewritten:

$$\begin{aligned} \sum_{i=1}^N \sum_{j=1}^N |y_i - y_j| &= \sum_{y_j \geq y_i} (y_j - y_i) + \sum_{y_j < y_i} (y_i - y_j) \\ &= 2 \sum_{y_j \geq y_i} (y_j - y_i) \\ &= 2 \sum_{i=1}^N \sum_{j=i}^N (y_j - y_i) \\ \frac{1}{2N^2} \sum_{i=1}^N \sum_{j=1}^N |y_i - y_j| &= \frac{1}{N^2} \sum_{i=1}^N \sum_{j=i}^N (y_j - y_i) \end{aligned}$$

We now transform this double sum into a double integral, by noting that if $x = i/N$, $\mathcal{F}^{-1}(x) = y_i$:

$$\begin{aligned} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=i}^N (y_j - y_i) &= \int_0^1 \int_x^1 (\mathcal{F}^{-1}(t) - \mathcal{F}^{-1}(x)) \, dt \, dx \\ &= \int_0^1 \left[\int_x^1 (\mathcal{F}^{-1}(t) - \mathcal{F}^{-1}(x)) \, dt \right] dx \\ &= \int_0^1 \left[\int_x^1 \mathcal{F}^{-1}(t) \, dt - (1-x)\mathcal{F}^{-1}(x) \right] dx \end{aligned}$$

Note that

$$\int_0^1 \mathcal{F}^{-1}(x) \, dx = \bar{y}$$

and

$$\int_x^1 \mathcal{F}^{-1}(t) dt = \int_0^1 \mathcal{F}^{-1}(t) dt - \int_0^x \mathcal{F}^{-1}(t) dt = \bar{y} - \mathcal{L}(x)$$

Therefore

$$\begin{aligned} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=i}^N (y_j - y_i) &= \int_0^1 \left[\int_x^1 \mathcal{F}^{-1}(t) dt - (1-x)\mathcal{F}^{-1}(x) \right] dx \\ &= \int_0^1 [\bar{y} - \mathcal{L}(x) - \mathcal{F}^{-1}(x) + x\mathcal{F}^{-1}(x)] dx \\ &= \int_0^1 [x\mathcal{F}^{-1}(x) - \mathcal{L}(x)] dx \end{aligned}$$

We integrate the first term under the integral by parts, noting that a primitive of $\mathcal{F}^{-1}$ is $\bar{y}\mathcal{L}$ according to definition (4.1):

$$\begin{aligned} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=i}^N (y_j - y_i) &= \int_0^1 [x\mathcal{F}^{-1}(x) - \mathcal{L}(x)] dx \\ &= \int_0^1 x\mathcal{F}^{-1}(x) dx - \int_0^1 \mathcal{L}(x) dx \\ &= \left[x \cdot \bar{y}\mathcal{L}(x) \right]_0^1 - 2 \int_0^1 \mathcal{L}(x) dx \\ &= 1 - 2 \int_0^1 \mathcal{L}(x) dx \end{aligned}$$

which proves equality (4.39).

4.A.2 Normalization of the exposure index

The normalized exposure index is defined by

$$P = \frac{1}{p(1-p)} \sum_{k=1}^K \frac{N_k}{N} (p_k - p)^2 \quad (4.40)$$

We are interested in the extreme values of

$$f(\mathbf{p}) = \sum_{k=1}^K N_k (p_k - p)^2 \quad (4.41)$$

subject to

$$\sum_{k=1}^K N_k p_k = p \quad (4.42)$$

where $\mathbf{p} = (p_k)_{k \in \{1, \dots, K\}} \in [0; 1]^K$.

We introduce the Lagrangian:

$$\mathcal{L}(\mathbf{p}, \lambda) = \sum_{k=1}^K N_k (p_k - p)^2 - \lambda \left(\sum_{k=1}^K N_k p_k - p \right) \quad (4.43)$$

and we differentiate with respect to p_k :

$$\frac{\partial \mathcal{L}}{\partial p_k}(\mathbf{p}, \lambda) = 2N_k(p_k - p) - \lambda N_k \quad (4.44)$$

The derivative equals 0 if and only if $p_k = p + \lambda/2$. Plugging this equality in the constraint (4.42) shows that λ must equal zero, therefore $p_k \equiv p$ for all k . This case corresponds to the minimal value of the exposure index ($P = 0$). Since the Lagrangian has no other extremum, the maximum values are reached on the corners of the domain $[0; 1]^K$, *i.e.* if $p_k \in \{0; 1\}$ for all k . In that case, $p_k^2 = p_k$, therefore with the constraint (4.42):

$$f(\mathbf{p}) = \sum_{k=1}^K N_k (p_k - p)^2 = \sum_{k=1}^K N_k (p_k - 2 \cdot p \cdot p_k + p^2) = N \cdot p - 2 \cdot N \cdot p^2 + N \cdot p^2 = N \cdot p(1 - p) \quad (4.45)$$

Hence the normalization in equation (4.40).

4.A.3 Normalized exposure is equal to the exposure gap

The normalized exposure index is defined by

$$P = \frac{1}{p(1-p)} \sum_{k=1}^K \frac{N_k}{N} (p_k - p)^2$$

We first want to show that

$$P = 1 - \frac{\mu^B}{p}$$

where μ^B is the average exposure of group B to group A.

Denote $N^B = N(1 - p)$ the number of individuals in group B and $N_k^B = N_k(1 - p_k)$ the number of individuals from group B in cluster k . Then, the definition of μ_B is:

$$\mu^B = \frac{1}{N^B} \sum_{k=1}^K N_k^B p_k = \frac{1}{N(1 - p)} \sum_{k=1}^K N_k(1 - p_k)p_k = \sum_{k=1}^K \frac{N_k}{N} \cdot \frac{p_k(1 - p_k)}{1 - p}$$

Multiply the first two equations by $p(1 - p)$:

$$\begin{aligned} p(1 - p)P &= \sum_{k=1}^K \frac{N_k}{N} (p_k - p)^2 \\ p(1 - p)P &= p(1 - p) - (1 - p)\mu^B = p(1 - p) - \sum_{k=1}^K \frac{N_k}{N} p_k(1 - p_k) \end{aligned}$$

We need to show that

$$\sum_{k=1}^K \frac{N_k}{N} (p_k - p)^2 = p(1 - p) - \sum_{k=1}^K \frac{N_k}{N} p_k(1 - p_k)$$

i.e.

$$p(1 - p) = \sum_{k=1}^K \frac{N_k}{N} (p_k - p)^2 + \sum_{k=1}^K \frac{N_k}{N} p_k(1 - p_k)$$

We start from the right hand side of this last equality:

$$\begin{aligned} \sum_{k=1}^K \frac{N_k}{N} (p_k - p)^2 + \sum_{k=1}^K \frac{N_k}{N} p_k(1 - p_k) &= \sum_{k=1}^K \frac{N_k}{N} [(p_k - p)^2 + p_k(1 - p_k)] \\ &= \sum_{k=1}^K \frac{N_k}{N} [p_k^2 - 2p \cdot p_k + p^2 + p_k - p_k^2] \\ &= p^2 + (1 - 2p) \sum_{k=1}^K \frac{N_k}{N} p_k \\ &= p^2 + (1 - 2p) \cdot p \\ &= p(1 - p) \end{aligned}$$

because the sum of N_k/N is one and the sum of $(N_k/N)p_k$ is the weighted average of p_k , which is equal to p . Hence the intermediary result:

$$P = 1 - \frac{\mu^B}{p}$$

Introduce μ^A , the average exposure of the reference group to itself. The average exposure of the population to the reference group is equal to p . It is also equal to the weighted average of the exposures of groups A and B, *i.e.*:

$$p = p \cdot \mu^A + (1 - p) \cdot \mu^B$$

Therefore, if $p \neq 0$

$$\mu^A = 1 - \frac{1 - p}{p} \mu^B$$

which yields the final result:

$$\mu^A - \mu^B = 1 - \frac{1 - p}{p} \mu^B - \mu^B = 1 - \frac{\mu^B}{p} = P$$

4.A.4 Upper envelope of the Lorenz curve of lifetime social environments

As stated in section 4.3.2, "in the best case scenario, there is no difference in $\overline{\mu}_i$ between individuals from the reference group and between individuals from the group B, *i.e.* $\overline{\mu}_i = \overline{\mu}^g$ if $i \in g$ where g denotes group A or B: if SEM is optimal, it will eliminate differences within each group, but it cannot eliminate differences between the two groups."

The values of the exposures in each group can be derived from equation (4.24):

$$\begin{aligned} \overline{\mu}^A &= p(1 - \overline{P}) + \overline{P} \\ \overline{\mu}^B &= p(1 - \overline{P}) \end{aligned}$$

Since group B has the lowest average exposure to the reference group and it represents a share $1 - p$ of the population, the inverse of the cumulative distribution function is given by

$$\mathcal{F}^{*-1}(u) = \begin{cases} \mu^B = p(1 - \overline{P}) & \text{if } u < 1 - p \\ \mu^A = p(1 - \overline{P}) + \overline{P} & \text{otherwise} \end{cases} \quad (4.46)$$

The cdf is piecewise constant, therefore the Lorenz curve, which is defined by

$$\mathcal{L}^*(u) = \frac{1}{p} \int_0^u \mathcal{F}^{*-1}(t) dt \quad (4.47)$$

is continuous and piecewise linear: it is a spline of degree one, with a knot at $u = 1 - p$ (at the level of the discontinuity in $\mathcal{F}^{*-1}$). We always have $\mathcal{L}^*(0) = 0$ and $\mathcal{L}^*(1) = 1$, therefore we only need to determine $\mathcal{L}^*(1 - p)$ in order to draw the curve in **FIGURE 4.6**:

$$\mathcal{L}^*(1 - p) = \frac{1}{p} \int_0^{1-p} p(1 - \bar{P}) \, dt = (1 - p)(1 - \bar{P})$$

