



HAL
open science

Inequalities of access to higher education : three institutional factors

Pauline Charousset

► **To cite this version:**

Pauline Charousset. Inequalities of access to higher education : three institutional factors. Economics and Finance. Université Panthéon-Sorbonne - Paris I, 2021. English. NNT : 2021PA01E037 . tel-03507751

HAL Id: tel-03507751

<https://theses.hal.science/tel-03507751v1>

Submitted on 3 Jan 2022

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

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

Université Paris 1 Panthéon Sorbonne - École d'Économie de Paris
UFR d'Économie
Laboratoire de rattachement : Paris Jourdan Sciences Économiques (PJSE)

THÈSE

Pour l'obtention du titre de Docteur en Sciences Économiques
Présentée et soutenue publiquement
le 21 juin 2021 par
Pauline CHAROUSSET

Inégalités d'accès à l'enseignement supérieur : trois facteurs institutionnels

Sous la direction de Julien GRENET

Composition du jury :

- Rapporteurs :* Elise HULLERY, Professeure à l'Université Paris-Dauphine
Arnaud MAUREL, *Associate Professor*, département d'économie de Duke
- Examineurs :* Marc GURGAND, Directeur de recherche CNRS - École d'Économie de Paris
Camille TERRIER, *Assistant Professor*, département d'économie d'HEC Lausanne
- Directeur de thèse :* Julien GRENET, Directeur de recherche CNRS - École d'Économie de Paris

Université Paris 1 Panthéon Sorbonne - Paris School of Economics

UFR d'Économie

Laboratory: Paris Jourdan Sciences Économiques (PJSE)

Ph.D. THESIS

Submitted for the Degree of Doctor of Philosophy in Economics

Defended on June 21, 2021, by:

Pauline CHAROUSSET

Inequalities of Access to Higher Education: Three Institutional Factors

Thesis advisor: Julien GRENET

Jury:

Reviewers : Elise HUILLERY, Professor at University Paris-Dauphine
Arnaud MAUREL, Associate Professor in the Department of Economics at Duke University

Examinators : Marc GURGAND, Senior Researcher at CNRS - Paris School of Economics
Camille TERRIER, Assistant Professeur in the Department of Economics at HEC Lausanne

Advisor : Julien GRENET, Senior Researcher at CNRS - Paris School of Economics

Remerciements

Cette thèse a été préparée dans le cadre exceptionnel offert par l'École d'Économie de Paris. Peut-être parce que son projet de mettre la science économique au service de la société attire des personnalités à part, ce lieu m'a offert des opportunités de rencontres professionnelles d'une grande richesse. Je pense tout d'abord à mon directeur de thèse, Julien Grenet. Julien, tout au long de cette thèse, tes qualités de pédagogue m'ont aidée à tirer le meilleur parti de mon travail, et tes talents de chercheur m'ont donné l'assurance de pouvoir compter sur un jugement d'une qualité rare. Ton soutien actif pour débloquer l'accès aux données administratives a été déterminant, et si salutaire dans les moments de découragement. Enfin, je dois beaucoup aux opportunités que tu m'as offertes en m'associant à des projets aussi passionnants que porteurs, et en me proposant de travailler à l'Institut des Politiques Publiques dès ma première année de thèse.

J'ai également la chance de bénéficier d'un jury de thèse d'une grande qualité scientifique. Merci à Élise Huillery et Arnaud Maurel, pour avoir endossé leur rôle de rapporteurs bien au-delà de ce que le professionnalisme exigeait : vos remarques, extrêmement riches et pertinentes, ont été déterminantes lors de la finalisation de cette thèse, et à Marc Gurgand, qui a suivi cette thèse pas à pas, et Camille Terrier d'avoir accepté de rejoindre le jury. Mes remerciements vont également à mes coauteurs, Gabrielle Fack, pour ses retours toujours très fins et justes, YingHua He, *for an unforgettable summer of econometric discussions under the heatwave in Paris*, et Marion Monnet, pour m'avoir fait bénéficier de sa riche expérience professionnelle, et pour toutes les épreuves de la thèse surmontées de front.

Les équipes de la DEPP, du SIES et de l'Insee, en m'ouvrant l'accès à des données pour partie inédites, ont joué un rôle central dans cette thèse : merci à Cosima Bluntz, Pierre Boulet, Ceren Inan, Isabelle Kabla-Langlois, Alexandre Lebrere, Claire Margaria et Caroline Simonis Sueur. Des remerciements particulièrement appuyés pour Soazig Jolivet et ses bonnes ondes bretonnes, et

Fabienne Rosenwald, grâce à qui le projet d'appariement entre données administratives éducatives et salariales a pu se concrétiser, *in extremis*, dans le temps de ma thèse. Fabienne, ces derniers mois, notre collaboration sur la revue *Éducation & Formations* m'a offert, en plus de conditions matérielles idéales pour achever ma thèse, une expérience professionnelle dont je garderai un excellent souvenir : merci mille fois ! Enfin, je remercie la Chaire femmes et sciences de Dauphine, qui a financé ma quatrième année de thèse.

À rebours de tout ce que la théorie économique pourrait prédire, mes choix éducatifs ont toujours été déterminés par des rencontres ou des hasards. Si ce sont essentiellement des hasards qui m'ont fait connaître PSE, ce sont les rencontres que j'y ai faites qui m'ont donné envie d'y rester. Un immense merci donc, aux excellents pédagogues de PSE, dont l'implication, la clarté, et, il faut bien le dire, le talent, m'ont enthousiasmée pendant mes années de master et encouragée sur la voie de la recherche. Une mention spéciale pour Luc Behaghel, Antoine Bozio, Julien Grenet, Marc Gurgand et Thomas Piketty.

Les équipes administratives et pédagogiques de PSE et Paris 1 font preuve au quotidien d'un dévouement et d'une efficacité exceptionnels. Merci, entre autres, à Magali David, Mouez Fodha, Jean-Christophe Fouillé, Christelle Gauvrit, Sylvie Lambert, Anaïs Leveneur, Clément Michaud et Sylvain Riffé Stern. Des remerciements chaleureux pour Lucia Roxana Ban et Pauline Marmin, qui m'ont permis de bénéficier de conditions d'enseignement idéales, et ont ainsi directement contribué aux excellents souvenirs que je garderai des interactions avec les étudiants du master PPD. Un hommage particulièrement appuyé pour le travail de titan accompli au service informatique par Radja Aroquiaradja et Stéphane Brice : c'est toute l'institution – près de 150 doctorants et 140 chercheurs – qui repose sur vos épaules, respect ! Enfin, bien sûr, chapeau bas à Véronique Guillotin, qui a été une présence exceptionnelle tout au long de cette thèse. Véronique, avec toi, c'est un peu de PSE qui s'en va... Bonne route !

J'ai eu la chance de passer mes trois premières années de thèse dans les locaux de l'Institut des Politiques Publiques, entourée de chercheurs, doctorants et stagiaires excellents dans leur domaine et soucieux de contribuer à la convivialité de leur environnement de travail. Un grand merci donc à l'ensemble de l'équipe de l'IPP, qui porte, avec un succès plus que mérité – et grandissant ! – un

projet scientifique et citoyen fondamental (le « citoyen », c'est pour toi, Antoine !).

Aux pensionnaires successifs des bureaux R3-13 et R3-18/19, Nagui Bechichi, Benjamin Belrhomari, Mahdi Benjelloul, Cécile Bonneau, Paul Brandily, Paul Corbel, Quiterie Duco, Marion Leroutier, Youssef Souidi, et Georgia Thebault, merci pour toutes les pauses si salutaires. Un petit mot à part pour Youssef : nos chemins se sont croisés pendant les années de master pour ne plus vraiment se séparer, et c'est avec émotion que je me souviens de nos premiers débats sur la *zero conditional mean assumption* ou le sens de la triple négation en anglais. Puisqu'on en est aux années de master, *special thanks to Ariel Gruver for her helpful comments on the English version of my dissertation summary, and for her transatlantic support*. Cécile et Georgia : notre collaboration dans le cadre du projet sur la démocratisation des grandes écoles – dont l'échéance n'a pourtant pas surgi au moment le plus opportun ! – a été une véritable bouffée d'oxygène tout droit venue de vos contrées toulousaines et marseillaises. Et Quiterie, pour une belle amitié née du partage de nombreuses valeurs essentielles, comme le refus des palmiers en Bretagne.

Les longues heures passées sur l'ordinateur n'ont pas fait que du bien au physique, et je remercie Cecilia et Santiago pour m'avoir guidée sur le long chemin de la guérison. Pour le bien du corps, mais aussi de l'âme, la musique a été plus que jamais essentielle. Merci à tous ceux avec qui je partage cette passion depuis tant d'années maintenant : les moments passés avec vous font partie des plus chers à mes yeux. Une place particulière pour Patrice, qui a élargi mes horizons et fait de moi la musicienne, et donc un peu la personne, que je suis aujourd'hui. Enfin, merci à mes amis de tous bords, avec un petit mot pour ceux rencontrés sur les bancs du lycée : voilà que la trentaine nous guette déjà, et on dirait bien que la vie n'a pas tout à fait fini de nous apprendre qu'on ne s'est pas rapprochés par hasard !

Je remercie enfin ma famille pour son soutien indéfectible et sa présence chaleureuse. À mes parents et ma sœur Juliette : mon investissement des problématiques éducatives doit beaucoup aux enseignements tirés de vos parcours, qui démontrent chacun à leur manière la marge de manœuvre dont on dispose pour orienter sa vie. Et à ma grand-mère, qui, en me transmettant l'histoire de son lien brisé à l'école, a probablement posé la première pierre de cette thèse.

Summary

This dissertation assesses the contribution of three distinct institutional features of the schooling system to the inequalities of access to higher education: the selectivity of access to higher education, teacher feedback in Grade 12, and admission rules to selective post-secondary programs. I adopt a broad conception of inequalities, including those that stem from high school tracking as well as from socioeconomic background. The analysis is based on unique French administrative data and findings are in reference to the French context.

The first chapter shows that lowering the selectivity of access to high school graduation can mitigate certain schooling inequalities determined by the students' high school tracking. Exploiting the 2008–2011 reform of the French vocational high school track, I show that reducing time required to reach vocational high school graduation from four to three years substantially increased vocational students' access to high school graduation, higher education, and post-secondary technical programs' degrees. These educational effects are associated with improved school-to-work transitions and better labor market access at an early stage of the career.

The second chapter, written with Gabrielle Fack, Julien Grenet and YingHua He, investigates the equity of admissions to selective post-secondary programs (CPGE) and quantifies the amount of socio-economic stratification that can be attributed to these admission rules. We show that the CPGE programs give a preferential treatment to geographically-close applicants, and that ignoring applicants' demographic characteristics when making admissions decisions can moderately contribute to alter the socio-demographic composition of students selected into these programs.

The third chapter, written with Marion Monnet, explores the extent to which math teachers' feedback affects their students' performance and subsequent application choices in higher education. Basing our analysis on the feedback given by teachers to students in their Grade 12 academic

transcripts, we show that math teachers differentiate their vocabulary according to the students' gender. This differentiation reflects gender stereotypes in part, but also a teacher's effort to adapt her feedback to the profile she perceives. We find that students exposed to a teacher with stronger gender differentiation perform better at the end of high school examinations but end up making application choices similar to other students. We interpret these findings as evidence that performance is elastic to the efforts of the student, who reacts to feedback, unlike application choices, which appear less malleable.

Keywords: higher education, institutional barriers, application choices, admissions decisions

Résumé

Cette thèse évalue la contribution de trois caractéristiques institutionnelles du système scolaire aux inégalités d'accès et d'orientation dans le supérieur : la sélectivité de l'accès au supérieur, les appréciations scolaires des professeurs en classe de terminale, et les modalités de recrutement dans les formations sélectives du supérieur. Nous nous attachons à couvrir un large éventail d'inégalités éducatives, en étudiant aussi bien les inégalités liées à l'orientation des élèves après le collège, que les inégalités socio-spatiales ou de genre. L'analyse est effectuée à partir de données administratives pour partie inédites, dans le contexte français.

Nous montrons dans un premier chapitre que la sélectivité de l'accès au baccalauréat est un déterminant majeur des inégalités d'accès au supérieur liées à l'orientation post-troisième, à partir de l'évaluation de la réforme de la voie professionnelle de 2008–2011. En réduisant de quatre à trois ans la durée du cursus menant au baccalauréat pour les élèves de la voie professionnelle, la réforme a considérablement augmenté les taux d'accès au baccalauréat des élèves de la voie professionnelle, améliorant ainsi significativement leurs chances d'accès au supérieur ainsi qu'au diplôme du supérieur dans les formations professionnalisantes de niveau bac+1/2. Ces effets sur les trajectoires éducatives des élèves de la voie professionnelle sont associés à une meilleure insertion sur le marché du travail en début de carrière.

Le deuxième chapitre, coécrit avec Gabrielle Fack, Julien Grenet et YingHua He, se place du côté des formations sélectives du supérieur, et analyse les modalités de traitement des candidatures reçues par les classes préparatoires aux grandes écoles (CPGE), avec pour objectif d'analyser leur équité, ainsi que de quantifier leur contribution à la stratification socio-spatiale et de genre dans le supérieur. Nous montrons que les classes préparatoires aux grandes écoles réservent un traitement préférentiel aux élèves issus du lycée ou du département où la formation est située. Nos analyses suggèrent que si les CPGE sélectionnaient les élèves sur la base de leur seul profil scolaire, elles ne

parviendraient qu'imparfaitement à reproduire la composition socio-spatiale observée en leur sein.

Le troisième chapitre, coécrit avec Marion Monnet, exploite des méthodes d'analyse textuelle pour étudier la différenciation du vocabulaire employé par les professeurs de mathématiques dans les bulletins scolaires de terminale en fonction du sexe de l'élève. Nous montrons qu'à niveau de performances scolaires égal, les professeurs adaptent leur vocabulaire au sexe de l'élève, selon une différenciation qui reflète à la fois des stéréotypes de genre, mais également un souci d'adaptation au profil psychologique de l'élève. Les élèves exposés à des professeurs qui différencient plus fortement leur vocabulaire ont tendance à avoir de meilleurs résultats au baccalauréat, mais ne réalisent pas des choix d'orientation différents des autres élèves. Nous interprétons ces résultats comme le reflet d'une plus forte élasticité de la performance des élèves au travail fourni, lui-même sensible aux indications et pratiques pédagogiques du professeur, au contraire des choix d'orientation, qui semblent moins malléables.

Mots clés: enseignement supérieur, barrières institutionnelles, choix d'orientation, critères d'admission.

Introduction Générale

L'élargissement de l'accès à l'enseignement supérieur constitue un enjeu majeur de politique publique. Du point de vue économique, l'augmentation du niveau de qualification de la population active, en répondant aux besoins croissants du marché du travail en main-d'œuvre qualifiée, offre la réponse la plus crédible aux problèmes chroniques de chômage de masse et d'inégalités salariales dans les pays économiquement avancés. Du point de vue social, le diplôme du supérieur est aujourd'hui considéré comme l'un des principaux vecteurs de la mobilité sociale intergénérationnelle. À ce titre, l'accès juste et équitable au diplôme apparaît comme l'un des principaux garants de la méritocratie républicaine.

Malgré l'importance de ces enjeux, la performance des pays occidentaux en matière d'ouverture de l'accès au supérieur aux individus issus de milieux défavorisés, mais aussi de réduction des inégalités de réussite dans le supérieur, semble encore insuffisante. En France, alors que 75 % des enfants de cadres poursuivent des études supérieures en 2019, seuls 48 % des enfants d'ouvriers font de même (DEPP, 2020). Aux États-Unis, l'écart dans la probabilité d'accéder à l'enseignement supérieur entre un élève issu d'une famille à revenus élevés (quartile supérieur de la distribution des revenus) et un élève issu d'une famille à bas revenus (quartile inférieur) est de 51 points de pourcentage — un écart par ailleurs maintenu lorsqu'on s'intéresse à la probabilité d'obtenir un diplôme du supérieur (Bailey and Dynarski, 2011). Ces inégalités sont d'autant plus marquées que la sélectivité des formations considérées augmente. Deux exemples illustrent ce fait de façon éloquente. En France, les chances d'accès aux grandes écoles, symbole de l'élite scolaire française, étaient en 2016 plus de 10 fois supérieures pour les élèves issus de professions et catégories socioprofessionnelles (PCS) très favorisées par rapport aux élèves issus de PCS défavorisées (Bonneau et al., 2021). Aux États-Unis, sur la période 1997 à 2013, les enfants dont les parents ont des revenus appartenant au centile supérieur de la distribution des revenus avaient 77 fois plus de chance d'accéder à une

université de l'Ivy League que ceux dont les revenus des parents appartiennent au quintile inférieur de la distribution des revenus (Chetty et al., 2017).

Si les écarts de performance scolaire expliquent en partie les différences de taux d'accès à l'enseignement supérieur des élèves appartenant à différents groupes socio-démographiques, ils s'avèrent insuffisants pour rendre entièrement compte de celles-ci.

Ainsi, à niveau de performances scolaires égal, les élèves ne font pas les mêmes choix éducatifs en fonction du niveau de revenus de leurs parents (Hoxby and Turner, 2015), de leur genre (Brown and Corcoran, 1997) ou encore de leur origine géographique (Van Zanten, 2015; Guillois et al., 2014). La littérature en économie, en psychologie et en sciences sociales a identifié plusieurs facteurs susceptibles d'expliquer les inégalités d'accès et d'orientation dans le supérieur. En premier lieu, les contraintes financières peuvent exercer une influence sur les choix éducatifs des élèves du fait des coûts directs des études, notamment les frais d'inscription et de scolarité (Angrist, 1993; Bettinger, 2004; Fack and Grenet, 2015; Dynarski et al., 2018), mais également de leurs coûts indirects, et ce de façon d'autant plus marquée au-delà de l'âge d'instruction obligatoire où la poursuite d'études implique le renoncement à un salaire. Les individus issus de familles à bas niveau de revenus sont également plus exposés à des mécanismes psychologiques agissant comme des freins à l'investissement éducatif, tels que la peur de l'endettement (Field, 2009) ou encore une préférence trop marquée pour le présent (Oreopoulos, 2007). Le déterminisme social des choix éducatifs des élèves va cependant au-delà des facteurs liés aux contraintes financières. Il s'explique également par l'existence d'effets de pairs et de conformisme social (Akerlof and Kranton, 2000; Dalton et al., 2016), qui contribuent à détourner les élèves performants issus de milieux sociaux défavorisés des formations perçues comme réservées à une élite sociale. Ces facteurs sont renforcés par les effets de l'appartenance sociale sur l'estime de soi (Oyserman et al., 2006) et sur la perception de son propre niveau de performance scolaire (Guyon and Huillery, 2021), qui conduisent certains élèves à sous-estimer leurs chances de réussite dans ces formations. Dans l'enseignement supérieur, ces inégalités sociales sont exacerbées par l'accompagnement insuffisant qui est proposé aux élèves issus de milieux défavorisés au cours du processus d'orientation. Par exemple, l'étude de Castleman and Page (2015) a démontré que la complexité des démarches administratives nécessaires pour

soumettre sa candidature ou s'inscrire dans le supérieur peut s'avérer dissuasive pour certains profils d'élèves. Enfin, les inégalités d'accès à l'information conduisent les élèves les moins dotés en capital social à être plus exposés au non-recours aux aides financières à la poursuite d'études (Bettinger et al., 2012), et moins conscients des rendements de l'enseignement supérieur sur le marché du travail (Hastings et al., 2015; Oreopoulos and Dunn, 2013).

Les inégalités attribuables aux performances scolaires des élèves sont elles-mêmes pour partie le résultat de déterminismes sociaux. La recherche en psychologie et en économie a ainsi démontré que les stéréotypes associés aux compétences intellectuelles des différents groupes sociaux et ethniques ou encore en fonction du genre peuvent affecter la performance des élèves (Croizet et al., 2004; Cadinu et al., 2005; Nosek et al., 2009b). Par exemple, une étude de Stone et al. (1999) menée aux États-Unis révèle que selon qu'une tâche sportive est présentée comme évaluant les qualités athlétiques innées des élèves, ou au contraire leur intelligence sportive, les écarts de performance entre élèves *African American* et *White* s'inversent, les élèves *African American* ayant de meilleures performances que les élèves *White* dans le premier cas et des performances inférieures dans le second. En Inde, une étude de Hoff and Pandey (2006) a montré que le fait de révéler publiquement la caste d'enfants participant à une expérimentation suffit à l'émergence d'un écart de performance entre enfants issus de castes élevées et de castes basses lors de la résolution d'énigmes, alors même que cet écart est inexistant lorsque les enfants conservent leur anonymat.

L'influence de chacun de ces mécanismes sur les choix d'orientation des élèves peut être atténuée par des politiques et interventions ponctuelles, y compris à des stades avancés de la scolarité des élèves. Par exemple, les interventions de *role models* féminins dans le cadre scolaire en classe de terminale ou à l'université (Porter and Serra, 2020; Breda et al., 2020) s'avèrent efficaces pour limiter le poids des stéréotypes de genre des élèves dans leurs choix d'orientation. D'autres options, consistant par exemple à envoyer des courriers personnalisés sur certaines formations à des élèves présentant un fort potentiel scolaire mais susceptibles de réaliser des choix sous-optimaux du fait de leur appartenance sociale, ont donné des résultats positifs sur les taux d'inscription des élèves concernés dans ces formations (Hoxby and Turner, 2015; Dynarski et al., 2018). Ces interventions sont d'importance cruciale, car elles parviennent à limiter les déterminismes sociaux à l'œuvre dans

les choix éducatifs des élèves, y compris à des stades de la scolarité considérés comme tardifs. De ce fait, la multiplication de telles initiatives pourrait contribuer à modifier les parcours éducatifs d'un nombre non négligeable d'élèves. Cependant, l'histoire de l'expansion scolaire du 20^e siècle révèle que les changements majeurs de paradigme en matière d'accès à l'éducation n'ont été obtenus que lorsque le système éducatif dans son ensemble a été repensé de façon à ne pas enfermer les élèves dans des spécialisations trop précoces, et à garantir au plus grand nombre des opportunités éducatives similaires. Dans tous les pays occidentaux, la seconde moitié du 20^e siècle a ainsi été riche d'initiatives de politiques éducatives, qui, en restructurant le système éducatif, ont contribué à résorber les inégalités éducatives déterminées par l'origine sociale. Ces politiques ont essentiellement ciblé l'enseignement secondaire, mais elles ont affecté les trajectoires scolaires des élèves au-delà des niveaux visés, contribuant ainsi directement à l'élargissement de la base de recrutement du supérieur. Dans toute l'Europe, l'allongement de la durée légale de scolarisation obligatoire (Oreopoulos, 2007; Aakvik et al., 2010), puis les réformes du collège unique, qui ont reporté la spécialisation des élèves à un stade ultérieur de leur scolarité, ont été les principaux vecteurs de l'expansion scolaire de cette période. Ce sont ainsi deux facteurs qui ont été primordiaux : le degré de spécialisation des élèves et le stade de leur scolarité auquel elle s'opère, ainsi que la sélectivité de l'accès aux différents niveaux du système éducatif (Meghir and Palme, 2005; Guyon et al., 2012).

Le cas particulier de la France s'avère particulièrement instructif, puisqu'il illustre de façon marquante le contraste entre les périodes de refonte du système éducatif, en amont comme en aval de l'enseignement supérieur, qui ont été associées à des évolutions considérables dans les niveaux de qualification atteints par les élèves et les périodes de stabilité institutionnelle où la démocratisation du système est demeurée en berne. Entre 1960 et 2020, le vivier de recrutement de l'enseignement supérieur français s'est ainsi considérablement élargi, mais à un rythme très inégal au cours de la période. En 2018, les effectifs de l'enseignement supérieur français (2,6 millions d'étudiants) étaient huit fois supérieurs à ceux de 1960 (310 000 étudiants), reflétant en partie l'allongement de la durée moyenne des études supérieures, mais surtout la nette augmentation du taux d'accès aux études supérieures. Parmi les individus nés à la fin des années 1990, 60 % ont ainsi accédé à l'enseignement supérieur, contre 41 % parmi ceux qui sont nés au début des années 1970, soit

une hausse de 50 % en l'espace de moins de 30 ans. Deux facteurs institutionnels étroitement liés ont joué un rôle central dans cette évolution. Premièrement, la massification de l'enseignement supérieur a largement reflété celle de son principal vivier de recrutement, les nouveaux bacheliers, dont le nombre a crû significativement après l'instauration du baccalauréat professionnel en 1985. Deuxièmement, l'adaptation du système d'enseignement supérieur, avec le développement de filières courtes professionnalisantes, a offert des débouchés aux nouveaux entrants dans l'enseignement supérieur. Créés dès le début des années 1960, les instituts universitaires de technologie (IUT) et, plus encore, les sections de technicien supérieur (STS), ont connu une croissance exponentielle à partir de la fin des années 1980, les effectifs de ces dernières étant passés de 30 000 étudiants en 1970 à près de 240 000 en 2000. À partir des années 2000, cette dynamique s'est cependant infléchie : alors que le taux d'accès au supérieur a crû à un rythme d'environ 5 % par an entre 1960 et 2000, la croissance annuelle n'a été que de 0,6 % au cours de la période 2000–2010. Conséquence directe de la stagnation du taux d'accès au baccalauréat, le taux d'accès au supérieur n'a ainsi augmenté qu'à un rythme lent, les filières courtes du supérieur ne contribuant plus à l'essor qu'elles avaient insufflé depuis la fin des années 1980. En l'absence de changement institutionnel majeur, cette stagnation a perduré jusqu'au début des années 2010. Elle n'a pris fin qu'en 2011, sous l'impulsion de la rénovation de la voie professionnelle de 2008/2009 : pour la première fois en 15 ans, le taux de bacheliers dans une génération a dévié des 65 % pour atteindre 80 % d'une génération, soit l'objectif fixé pour l'an 2000 par la loi d'orientation sur l'éducation de la fin des années 1980.

L'objectif de cette thèse est d'identifier un certain nombre de facteurs institutionnels propres au système scolaire qui sont susceptibles de contribuer aux inégalités d'accès à l'enseignement supérieur. Parmi les facteurs envisagés, une attention particulière est portée à l'influence exercée par la structure des études secondaires (Chapitre 1), et les appréciations scolaires des professeurs (Chapitre 3) sur les choix d'orientation des élèves. Dans l'enseignement supérieur, nous nous intéressons aux modalités de recrutement des formations les plus sélectives (Chapitre 2). Avant de détailler la problématique commune aux trois chapitres, et de présenter le plan de cette thèse, nous exposons les intérêts de l'approche économique de l'évaluation des politiques éducatives, avec une attention particulière pour les questions centrales aujourd'hui envisagées par la discipline.

Une approche économique de l'évaluation des politiques éducatives

Dans la lignée des travaux de Gary Becker et Jacob Mincer, l'approche économique de l'éducation s'est d'abord penchée sur la question des rendements salariaux de l'éducation. Le cadre initialement développé propose d'interpréter l'éducation comme un investissement devant garantir, à moyen et long terme, une meilleure insertion sur le marché du travail et des salaires plus élevés. Pour que cet investissement soit rentable, il est nécessaire que ses rendements (notamment les gains salariaux) surpassent ses coûts. Cette approche implique l'existence d'un niveau d'études au-delà duquel il n'est plus pertinent, du point de vue des rendements salariaux privés de l'éducation, de poursuivre l'investissement éducatif, sous l'hypothèse de rendements marginaux décroissants et de coûts marginaux constants ou croissants avec le nombre d'années d'études. Au cours de la seconde moitié du 20^e siècle, les efforts empiriques pour tenter de déterminer la durée optimale des études se sont ainsi traduits, dans un premier temps, par l'étude des rendements salariaux de l'éducation à différents stades de la scolarité, avant d'être étendus à d'autres dimensions de la vie des individus, telles que la santé et la fertilité (Haveman and Wolfe, 1994), ou encore la criminalité (Lochner and Moretti, 2004).

La multiplication des sources de données renseignant à la fois le niveau d'éducation des individus, ainsi que leur niveau de vie à l'âge adulte a permis d'établir l'existence d'un lien statistique fort entre le niveau d'études et les salaires perçus au cours de la carrière. Cette relation statistique reflète cependant pour partie des différences intrinsèques entre les individus ayant suivi des études plus ou moins longues, qui sont susceptibles d'expliquer leurs performances sur le marché du travail indépendamment des effets de l'éducation sur leur carrière. Pour résoudre le défi méthodologique posé par la corrélation entre ces facteurs, souvent non observés par les économistes, et les choix éducatifs des individus ainsi que leur niveau de vie à l'âge adulte, la littérature en économie de l'éducation a largement exploité les changements opérés dans le système éducatif au cours de la seconde moitié du 20^e siècle comme sources de variation exogène permettant d'identifier des effets causaux. Les politiques d'expansion scolaire ont, dans certains cas, présenté l'intérêt de

toucher des publics similaires de façon différentielle au cours du temps, le plus souvent du fait de la régionalisation de la mise en œuvre des réformes (Gurgand and Maurin, 2006). Par exemple, en Suède, la réforme qui a étendu, dans les années 1950, le tronc commun du système scolaire suédois de 6 à 9 ans, n'a d'abord concerné qu'un tiers des communes, avant d'être généralisée aux deux tiers restants. Ce type de configuration est propice à l'identification causale des effets de l'éducation, parce qu'il génère des variations dans le niveau d'études atteint par des individus qui ne diffèrent, en moyenne, que selon une dimension, connue de l'économiste (par exemple la commune de résidence des élèves), et dont les effets directs potentiels sur les rendements salariaux peuvent être pris en compte. En exploitant ces situations, une abondante littérature consacrée aux rendements de l'éducation a identifié des effets positifs de l'enseignement secondaire sur les salaires, mais qui varient de façon importante en fonction du contexte et de la population étudiés, avec une hausse moyenne des salaires comprise entre 5 % et 15 % pour chaque année supplémentaire d'études secondaires (Angrist and Krueger, 1991; Harmon and Walker, 1995; Oreopoulos, 2007; Meghir and Palme, 2005; Duflo, 2001; Aakvik et al., 2010). Au-delà de l'âge d'instruction obligatoire, les effets de l'investissement scolaire sur les niveaux de vie sont moins clairement établis, particulièrement pour les élèves ayant suivi des études professionnelles dans le secondaire, qui ne bénéficient pas, en moyenne, d'un allongement de leurs études secondaires (Hall, 2016; Zilic, 2018; Bertrand et al., 2019; Oosterbeek and Webbink, 2007), ou encore pour ceux à la marge d'accéder à l'enseignement supérieur. Pour ces derniers, si une partie de la littérature conclut à des effets positifs de l'enseignement supérieur sur les salaires (Maurin and McNally, 2008; Zimmerman, 2014; Carneiro et al., 2011), les rendements salariaux associés semblent particulièrement hétérogènes. Les récents développements de la littérature suggèrent en particulier que l'accès au diplôme du supérieur ne bénéficie pas à tous les profils d'élèves, les rendements associés étant surestimés en l'absence de prise en compte des différentiels cognitifs entre élèves (Heckman et al., 2018; Ashworth et al., 2017).

Au-delà des facteurs individuels, l'hétérogénéité des rendements de l'éducation dans l'enseignement supérieur s'explique par le fait de considérer un stade avancé de la scolarité, et par la nature intrinsèquement différente de l'investissement réalisé dans le supérieur par rapport au secondaire. La différenciation institutionnelle et le degré de spécialisation propres au supérieur complexifient

l'étude des rendements salariaux de l'éducation, puisqu'à ce stade de la scolarité, l'investissement éducatif ne se résume plus seulement à un arbitrage entre le fait de poursuivre des études ou rejoindre le marché du travail. Outre la détermination de la quantité à investir, la question de la nature de l'investissement à réaliser devient centrale. Pour les décideurs publics, l'objectif n'est donc pas seulement de déterminer la pertinence d'inciter les élèves à poursuivre des études supérieures, mais également d'assurer la meilleure adéquation entre étudiants, d'une part, et formations du supérieur, d'autre part. Or, dans de nombreux pays économiquement avancés, les forts taux d'échec et les nombreuses réorientations des étudiants dans le premier cycle universitaire ont laissé craindre l'existence d'une inefficience allocative à l'entrée dans l'enseignement supérieur. Cette inefficience se manifesterait selon plusieurs dimensions, certains étudiants réalisant des choix qui ne maximiseraient pas leurs chances de réussite dans le supérieur, tandis que d'autres au contraire s'avèreraient trop peu ambitieux (Smith et al., 2013) ou ne choisiraient pas les filières présentant les meilleurs rendements sur le marché du travail (Altonji et al., 2012).

Si les difficultés méthodologiques ont jusqu'à ce jour empêché la production d'études qui quantifieraient de façon systématique l'ampleur de cette inefficience allocative à l'œuvre dans l'enseignement supérieur, de nombreux travaux en ont en revanche exploré les causes. Les facteurs financiers, informationnels, ou psychologiques, mentionnés plus haut comme des vecteurs d'inégalités dans l'accès à l'enseignement supérieur, semblent également affecter le degré d'optimalité des choix des élèves lors de l'orientation dans le supérieur. Parce qu'elle concerne majoritairement les élèves issus de milieux sociaux défavorisés ou de territoires où l'on accède à une information de moins bonne qualité sur les formations du supérieur, l'inefficience allocative serait ainsi un vecteur important d'inégalités d'accès au diplôme du supérieur à court terme, mais également d'inégalités salariales à plus long terme. La principale réponse proposée dans la littérature pour résoudre ces difficultés a été l'accompagnement des élèves. Une question, relativement sous-étudiée, est celle de la contribution des formations du supérieur à l'amélioration de l'adéquation entre étudiants et formations, d'une part, et à la limitation des inégalités socio-démographiques, d'autre part. Dans les formations sélectives, les critères d'admission pourraient en effet viser en premier lieu à assurer un recrutement qui maximise les chances de réussite des étudiants, corrigeant ainsi les éventuelles asymétries

d'information entre formations et étudiants sur les chances de réussite. Par ailleurs, ils pourraient chercher à œuvrer en faveur de l'égalisation des chances d'accès entre groupes sociaux ou encore de la diversité au sein de la formation en réservant un traitement préférentiel à certains profils d'élèves, sans que cet objectif soit nécessairement compatible avec celui de corriger les asymétries informationnelles.

Ce possible arbitrage entre efficience allocative et réduction des inégalités socio-démographiques d'accès aux formations sélectives du supérieur a été identifié pour la première fois dans les contextes de discrimination positive. Initialement mis en place pour corriger les inégalités d'accès à l'enseignement supérieur, ce type de politique vise à favoriser l'accès de certains groupes ethniques ou sociaux aux formations sélectives en réservant aux individus issus de ces groupes un traitement préférentiel. C'est par exemple le cas de l'*affirmative action* aux Etats-Unis, d'abord fondée sur critères ethniques puis remplacée par la version *race neutral*, avec par exemple le programme « *Top 10 percent rule* » mis en place au Texas en 1997 (Kapoor, 2015). En France, ce sont les inégalités sociales d'accès aux études supérieures que ces politiques ont cherché à corriger, avec le dispositif national meilleurs bacheliers, ou encore des initiatives de certains établissements sélectifs du supérieur, comme Sciences Po à travers les Conventions Éducation Prioritaire. Le sujet de la discrimination positive reste cependant très controversé dans la littérature économique (Holzer and Neumark, 2000; Arcidiacono and Lovenheim, 2016). Plusieurs études récentes ont montré que les politiques de discrimination positive pouvaient avoir des effets pervers pour les individus qu'elles sont supposées avantager (Kojima, 2012; Arcidiacono et al., 2015; Arcidiacono and Lovenheim, 2016), notamment en favorisant une forme particulière d'inefficience allocative, le « *mismatch* » (Sander, 2004) : les politiques de discrimination positive pourraient pénaliser les supposés bénéficiaires car ces derniers se retrouveraient affectés à des formations qui ne maximisent pas nécessairement leurs chances de succès. En dehors des contextes de discrimination positive, il existe peu d'études documentant les critères retenus par les formations sélectives pour recruter leurs étudiants, à l'exception des travaux qui se sont penchés sur l'étude des incitations que peuvent avoir les formations sélectives du supérieur à choisir stratégiquement leurs étudiants pour répondre à des contraintes liées à leur financement, notamment aux États-Unis (Andrews and Stange, 2019; Fu, 2014; Groen and White,

2004).

Problématique et plan de thèse

Cette thèse s'articule en trois chapitres indépendants, dont chacun explore la contribution d'un facteur institutionnel spécifique aux inégalités socio-démographiques et spatiales d'accès et de réussite dans l'enseignement supérieur, dans le contexte de la France des années 2000 à nos jours. Chaque chapitre apporte ainsi un éclairage particulier à la question centrale suivante : dans quelle mesure les facteurs institutionnels déterminent-ils les opportunités éducatives offertes aux élèves issus de différents groupes socio-démographiques lors de la transition entre enseignement secondaire et supérieur ? Le premier chapitre explore la manière dont la structure des études secondaires affecte les chances d'accès au baccalauréat, et ainsi à l'enseignement supérieur. Le deuxième chapitre analyse l'équité sociale, géographique et de genre des modalités d'admission dans les formations les plus sélectives de l'enseignement supérieur : les classes préparatoires aux grandes écoles. Le troisième chapitre, enfin, s'intéresse au rôle joué par les appréciations scolaires des professeurs de terminale dans les performances scolaires et la formation des choix d'orientation post-bac de leurs élèves.

Chapitre 1 : Lever les barrières institutionnelles à l'accès au supérieur pour les élèves de la voie professionnelle : effets sur les trajectoires scolaires et les niveaux de vie en début de carrière

Le premier chapitre de cette thèse explore le rôle des barrières institutionnelles à l'accès au supérieur pour les élèves de la voie professionnelle, en analysant les effets de la rénovation de la voie professionnelle mise en œuvre en France entre 2008 et 2011. Nos analyses ont été menées sur la base de données administratives éducatives produites par les services statistiques des ministères de l'Éducation Nationale, de la Jeunesse et des Sports (MENJS-DEPP) et de l'Enseignement Supérieur, de la Recherche et de l'Innovation (MESRI-SIES), ainsi que de données administratives salariales produites par l'Insee.

La rénovation de la voie professionnelle a consisté en une refonte majeure de la structure de l'enseignement professionnel dispensé en lycée. Avant la réforme, le cursus menant au baccalauréat professionnel s'étendait sur quatre années d'études, dont deux années professionnalisantes, qui permettaient l'obtention d'un diplôme professionnel – le brevet d'études professionnelles (BEP) – suivies de deux années plus généralistes – la première et la terminale professionnelles – qui préparaient les élèves aux examens du baccalauréat professionnel. Sur la base des données issues des dispositifs SYSCA, SISE et OCEAN (MENJS-MESRI), qui recensent annuellement l'ensemble des élèves et étudiants inscrits dans l'enseignement secondaire et supérieur français, nous montrons qu'en uniformisant le cursus menant au baccalauréat professionnel, et en le calquant sur le modèle en trois ans de la voie générale et technologique (seconde/première/terminale), la réforme de la voie professionnelle a permis d'augmenter sensiblement, après plus d'une décennie de stagnation, les taux d'accès au baccalauréat et à l'enseignement supérieur des cohortes entrées au lycée après 2008. Ce faisant, elle a réduit les inégalités d'accès à l'enseignement supérieur entre les élèves de la voie générale et technologique et les élèves de la voie professionnelle. Dans l'enseignement supérieur, les effets de la réforme ne se sont pas cantonnés à une augmentation symbolique des taux d'accès aux études supérieures : grâce notamment à l'offre de formations courtes professionnalisantes, en particulier les sections de techniciens supérieurs, la réforme a accru les chances d'accès à un diplôme du supérieur pour les élèves de la voie professionnelle.

La réforme de la voie professionnelle a donc constitué, au cours de la décennie passée, le principal levier de démocratisation de l'accès à l'enseignement supérieur, mais aussi (bien que dans une moindre mesure) au diplôme du supérieur. Pour déterminer si elle a pleinement porté ses fruits, ce chapitre étudie dans un second temps si cette réforme est parvenue à réduire les inégalités salariales déterminées par l'orientation post-troisième. L'obtention des autorisations juridiques pour relier les données administratives sur les trajectoires des élèves à des données individuelles sur les salaires extraites des déclarations annuelles de données sociales (DADS/DSN) permet d'apporter des éléments de réponse à cette question. La réforme de la voie professionnelle a eu des effets modérés mais positifs sur la stabilité de l'emploi et les salaires de début de carrière des élèves de la voie professionnelle. Neuf ans après la troisième, la part des élèves de la voie professionnelle

titulaires d'un contrat de travail à durée indéterminée a augmenté de 2,2 points de pourcentage du fait de la réforme, passant ainsi de 50,7 % à 52,9 %. Bien que la réforme n'ait eu qu'un impact limité sur le salaire moyen des élèves qui ont rejoint la voie professionnelle après la classe de troisième, certains groupes d'élèves de la voie professionnelle ont bénéficié d'augmentations non négligeables de leur salaire annuel, avec des hausses de 5,4 % pour les élèves de sexe masculin et de 4,5 % pour les plus performants scolairement (identifiés comme les 50 % des élèves qui ont obtenu les meilleurs résultats au brevet).

Ce chapitre contribue à deux sous-ensembles distincts de la littérature. Premièrement, il s'inscrit dans la littérature des barrières institutionnelles à l'éducation, qui comprend l'ensemble des travaux ayant étudié les réformes du collège unique et de l'allongement de l'âge d'instruction obligatoire, et démontre que ces barrières sont toujours à l'œuvre à un stade de la scolarité plus tardif que celui auquel elles ont été préalablement étudiées. À notre connaissance, seule l'étude de Maurin and McNally (2008) a formellement montré le rôle de la sélectivité de l'accès au diplôme de fin d'études secondaires (le baccalauréat dans le contexte français) comme barrière à l'entrée du supérieur.

Deuxièmement, ce chapitre apporte de nouveaux éclairages sur les rendements des études professionnelles et supérieures pour un groupe d'élèves sous-étudié au regard de son poids dans la population et de l'importance des enjeux de politiques publiques qu'il concentre : les élèves de la voie professionnelle. Pour les élèves qui auraient obtenu, dans l'ancien système, le baccalauréat, nous confirmons le résultat dominant de la littérature, selon lequel les élèves de la voie professionnelle ne bénéficient pas, du point de vue des rendements salariaux, d'un allongement de leurs études secondaires. Pour les autres élèves de la voie professionnelle, en revanche, nous montrons que l'exposition au cursus menant au baccalauréat en trois ans se traduit par une hausse de la probabilité d'être titulaire d'un contrat de travail à durée indéterminée en début de carrière.

Chapitre 2 : Quelle équité des règles d'admission dans les formations les plus sélectives du supérieur ?

Le deuxième chapitre de cette thèse propose une méthode d'analyse du traitement des candidatures reçues par les classes préparatoires aux grandes écoles (CPGE), qui sont les formations de niveau

bac+1 les plus sélectives, et évalue son impact sur la stratification sociale, géographique, et de genre observée à l'entrée dans le supérieur. Les analyses réalisées dans ce chapitre s'appuient sur des données issues du dispositif Admission Post-Bac (APB), qui centralisait jusqu'en 2017 la procédure d'admission des élèves dans les formations d'enseignement supérieur de niveau bac+1¹.

La contribution de ce chapitre à la compréhension des modalités d'admission dans les formations sélectives du supérieur est double. D'un point de vue méthodologique, nous montrons comment exploiter des méthodes de *reverse engineering* pour recouvrer les règles de classement des candidats par les CPGE à partir de l'observation de ces classements ainsi que de l'ensemble des informations utilisées par les CPGE pour les générer. Dans un second temps, nous utilisons les règles de classement des CPGE pour mesurer la contribution de leurs préférences à la stratification observée à l'entrée dans l'enseignement supérieur, en simulant les effets d'une politique qui imposerait aux CPGE de recruter les candidats sur la base de leur seul profil scolaire, sans avoir accès aux informations relatives à leur genre, origine sociale, lycée et département d'origine.

L'analyse des préférences des CPGE révèle qu'à profil scolaire égal, ces formations ne réservent pas le même traitement aux candidats en fonction de leur lycée ou département d'origine, et, dans une moindre mesure, de leur genre ou de leur origine sociale. Pour quantifier ce traitement différencié des candidatures, nous faisons l'hypothèse que les candidats sont classés à partir d'un score, calculé sur la base de la moyenne pondérée de leurs notes au contrôle continu ou aux épreuves anticipées du baccalauréat, qui peut être affiné à partir de bonus ou de malus appliqués aux caractéristiques socio-démographiques des candidats. Cette approche permet de mettre en évidence le traitement préférentiel dont bénéficient les élèves les moins distants géographiquement de la formation, et ce, de façon d'autant plus marquée que l'échelon géographique retenu est fin. Les résultats de l'étude montrent ainsi que 25 % des CPGE scientifiques réservent aux candidats qui ont obtenu leur baccalauréat dans le lycée où la formation est située un bonus équivalent à une augmentation de 30 % d'un écart-type de chacune des notes retenues par les formations pour classer les candidats, et à un avantage équivalent à 16 % d'un écart-type pour les candidats originaires du département où est située la formation. Pour les CPGE commerciales et littéraires, l'avantage conféré aux élèves originaires du lycée par les 25 % de CPGE ayant le plus fort degré de préférence

¹Ce chapitre a été coécrit avec Gabrielle Fack, Julien Grenet et YingHua He.

pour la proximité géographique est de l'ordre de 16 à 17 % d'un écart-type, et celui conféré aux élèves du département est compris entre 7 et 10 % d'un écart-type. Les bonifications ou pénalités accordées en fonction du genre ou de l'origine sociale sont de plus faible ampleur : l'avantage ou le désavantage donné aux candidats sur cette base est inférieur à 4 % d'un écart-type, en valeur absolue, pour la moitié des CPGE (tous types confondus).

Nos simulations suggèrent qu'interdire aux CPGE l'utilisation des informations relatives aux caractéristiques socio-démographiques des candidats aurait, en moyenne, peu d'impact sur les caractéristiques socio-démographiques et scolaires des candidats admis, à l'exception de la dimension géographique. Notre spécification principale suggère ainsi que la mise en place d'une telle politique ferait passer la proportion d'élèves originaires du lycée où est située la CPGE considérée de 9,1 % à 8,4 % parmi les candidats admis à l'issue de la première phase d'affectation d'APB.

En étudiant le contexte français, qui, à la différence des autres contextes préalablement étudiés (notamment les États-Unis, le Brésil et l'Inde) ne met pas en œuvre de politique de discrimination positive à grande échelle, nous mettons ainsi à jour l'existence d'une préférence pour la proximité géographique des étudiants. À notre connaissance, cette dernière n'avait jusqu'alors été démontrée que dans le contexte des États-Unis, où certaines formations financées publiquement sont incitées à recruter préférentiellement des étudiants originaires de l'État où est implantée la formation, mais augmentent le recrutement des étudiants originaires d'autres États lorsque la part des subventions étatiques dans leur financement diminue (Groen and White, 2004; Jaquette and Curs, 2015).

Chapitre 3 : Différenciation genrée des appréciations scolaires des professeurs, performances scolaires et choix d'orientation des élèves.

Le troisième chapitre de cette thèse analyse la différenciation des appréciations des professeurs de mathématiques en classe de terminale selon le sexe des élèves, avant d'explorer son rôle sur la performance scolaire et les choix d'orientation post-bac. Les analyses menées dans ce chapitre reposent sur les données d'APB (2012–2017), qui recensent l'ensemble des informations contenues dans les dossiers de candidature des élèves à l'entrée dans le supérieur, y compris leurs bulletins de

terminale².

À partir de l'analyse textuelle des appréciations figurant dans les bulletins scolaires, nous montrons qu'à niveau de performances scolaires égal, les professeurs de mathématiques adaptent le vocabulaire employé au sexe de l'élève. De façon à mettre en lumière cette différenciation genrée, nous utilisons un modèle de *machine learning* pour prédire le sexe des élèves à partir du vocabulaire figurant dans les appréciations formulées par leurs professeurs. Alors qu'un vocabulaire neutre conduirait à ne prédire correctement le sexe que de 50 % des élèves (soit le taux de prédictions correctes qu'on obtiendrait en s'en remettant au hasard), le vocabulaire présent dans les appréciations scolaires permet de prédire correctement le sexe de plus de 60 % des élèves, ce qui correspond à un taux de prédictions correctes marginalement inférieur à celui qu'on obtient lorsqu'on cherche à prédire si l'élève appartient à la moitié inférieure ou supérieure de sa classe en termes de performances scolaires.

Nous montrons que cette différenciation genrée du vocabulaire employé reflète à la fois des stéréotypes de genre, mais également un souci d'adaptation au profil psychologique de l'élève. Ainsi, les professeurs dont le vocabulaire est plus fortement différencié en fonction du sexe des élèves ont tendance à insister sur les aptitudes innées des élèves de sexe masculin, mais également à encourager les efforts des élèves de sexe féminin, dont la littérature a montré qu'elles sont plus sujettes au stress (Hill et al., 2016) et à la sous-estimation (Skaalvik and Skaalvik, 2004) dans les matières scientifiques.

Nous faisons l'hypothèse que la différenciation genrée du vocabulaire employé par le professeur dans les appréciations scolaires reflète de manière plus générale une différenciation dans la nature de ses interactions avec les élèves de sexe masculin et féminin tout au long de l'année scolaire, et explorons son impact sur la réussite et les aspirations éducatives des élèves. Pour ce faire, nous construisons un indicateur synthétique de différenciation genrée du vocabulaire, calculé pour chaque professeur à partir du taux d'observations correctement classifiées par le modèle : plus un professeur différencie son vocabulaire en fonction du sexe de ses élèves, plus la part d'observations dont le sexe est correctement prédit sera importante, et plus élevée sera la valeur de l'indicateur. Nous montrons que les élèves exposés à des professeurs qui différencient plus fortement leur vocabulaire en fonction

²Ce chapitre de thèse a été coécrit avec Marion Monnet.

du sexe des élèves ont tendance à avoir de meilleurs résultats au baccalauréat, mais ne réalisent pas des choix d'orientation différents des autres élèves. Nous interprétons ces résultats comme le reflet d'une plus forte élasticité de la performance des élèves au travail fourni, lui-même sensible aux indications et pratiques pédagogiques du professeur, au contraire des choix d'orientation, qui semblent moins malléables.

Ce chapitre contribue à la grandissante littérature reposant sur des méthodes d'analyse textuelle pour étudier les appréciations des professeurs. Si les bulletins scolaires ont depuis longtemps été envisagés comme un objet d'étude en sciences sociales, notamment avec les travaux pionniers de Bourdieu and Saint Martin (1975) qui ont illustré la différenciation des appréciations scolaires en fonction de la profession des parents, les nouvelles méthodes d'analyse textuelle offrent la possibilité de réaliser des analyses systématiques portant sur des cohortes entières d'élèves. À notre connaissance, notre étude est la première à réaliser ce type d'analyses sur des données exhaustives couvrant l'ensemble des élèves de terminale scientifique entre 2012 et 2017.

Ce chapitre contribue également à la littérature sur les effets de l'évaluation par un tiers en position d'autorité (ici, le professeur) du travail fourni sur la performance et la représentation de son propre niveau de performance. Nos résultats corroborent ceux de la littérature en psychologie, qui a démontré que les évaluations insistant sur les efforts fournis (*effort-based feedback*) favorisent l'investissement et la progression des élèves, au contraire des commentaires relatifs à la nature des individus (*person-based feedback*).

Contents

Acknowledgements	v
Summary	ix
General Introduction	1
1 Does Time to Graduation Affect Wages? Evidence from the French Vocational Track Reform	21
Introduction	23
1.1 Institutional Background	28
1.1.1 The French Schooling System	28
1.1.2 The 2008 Vocational Track Reform	30
1.2 Empirical Strategy	32
1.2.1 An IV Approach	32
1.2.2 Identifying Assumptions	34
1.3 Data and Summary Statistics	36
1.3.1 Data Sources	36
1.3.2 Sample Definition and Summary Statistics	38
1.4 Results	39
1.4.1 Graphical Evidence on the Impact of the Reform	39
1.4.2 Validity of the IV Strategy	41
1.4.3 Main Results	43
1.4.4 Heterogeneity Analysis	44
1.4.5 Robustness Checks	47
1.5 Discussion	48
1.5.1 Potential Mechanisms	48
1.5.2 Impact of the Reform by Pre-Reform Student Type	50
1.5.3 Exploring the Gender Heterogeneity in the Labor Market Effects of the Reform	51
1.6 Conclusion	53
Appendices	74
2 How Do Elite Post-Secondary Programs Select their Students?	107
Introduction	108
2.1 Institutional Background	114
2.2 Data and Summary Statistics	117
2.2.1 Data Sources	117
2.2.2 Student Characteristics	119

2.2.3	Sample Restrictions and Summary Statistics	119
2.3	How Each Program Ranks Its Applicants	121
2.3.1	A Conditional Logit Model	121
2.3.2	Penalized Conditional Logistic Regression	122
2.4	Programs' Preferences and Rankings of Students	123
2.4.1	Programs' Preferences	123
2.4.2	Programs' Rankings of Students	126
2.4.3	Controlling for Predicted <i>Baccalauréat</i> Grades	127
2.5	Counterfactual Simulations	128
2.5.1	Impact of a Ban on Applicants' Characteristics	129
2.5.2	Heterogeneous Effects of the Ban on Applicants' Characteristics across Different Types of CPGE	130
2.6	Discussion	131
2.6.1	Would the Policy Ban on Applicants' Characteristics Be Costly for CPGE Programs?	132
2.6.2	Effect of the Policy Ban when Students React to It	132
2.7	Conclusion	133

Appendices **153**

3 Gendered Teacher Feedback, Students' Math Performance and Enrollment Outcomes: A Text Mining Approach **175**

Introduction	177
3.1 Institutional Background	181
3.1.1 The French Secondary Education System	181
3.1.2 Tertiary Application and Enrollment	183
3.2 Data and Summary Statistics	184
3.2.1 Data Sources	184
3.2.2 Summary Statistics	186
3.3 Empirical Strategy	187
3.3.1 Measuring Teachers' Gender Differentiated Vocabulary (GDV)	187
3.3.2 Identification Strategy	191
3.4 Math Gender-Differentiated Feedback	191
3.4.1 Descriptive Statistics on Math Feedback by Gender	191
3.4.2 Math Teachers' Gender Differentiated Vocabulary (GDV)	193
3.5 Impact of Teachers Gender-Differentiated Vocabulary on Students' Outcomes	196
3.5.1 Validity of the Empirical Strategy	196
3.5.1.1 Exogeneity Assumption	196
3.5.1.2 Reverse Causality	197
3.5.2 Impact of Teacher GDV on Performance and Enrollment	198
3.6 Mechanisms	200
3.6.1 Teacher Grading Bias	200
3.6.2 Teacher Quality	201
3.6.3 Teacher's Type of Vocabulary Used	202
3.7 Conclusion	203
Figures	204
Tables	217

Appendices	226
<i>Bibliography</i>	265
<i>List of Tables</i>	269
<i>List of Figures</i>	273

Chapter 1

Does Time to Graduation Affect Wages? Evidence from the French Vocational Track Reform *

In 2008, a major reform of the high school vocational track was implemented in France to increase the average schooling attainment level of vocational track students. The reform reduced the time to vocational high school graduation by replacing the two-stage four-year vocational *baccalauréat* curriculum by an integrated three-year curriculum. Exploiting variation in exposure to the reform within middle schools over time, I show that the reform considerably improved the educational attainment level of vocational track students, as measured by their access to academic and vocational high school credentials as well as to higher education. Using administrative data covering the universe of students enrolled in Grade 9 between 2006 and 2009 and of wage earners between 2016 and 2018, I estimate that the reform increased vocational track students' probability of completing the *baccalauréat* by 10.0 percentage points, from a baseline of 26.5 percent. Moreover, the reform increased vocational students' rate of access to higher education by 3.1 percentage points, from a baseline of 10.9 percent. The reform led to moderate improvements in labor market outcomes, with

*I am grateful to the staff at the French Ministry of Education (Ministère de l'Éducation Nationale, Direction de l'Évaluation, de la Prospective et de la Performance) and the Insee for their invaluable assistance in accessing the data. This paper greatly benefited from discussions and helpful comments from Clément de Chaisemartin, Gabrielle Fack, Marc Gurgand, Élise Huillery and Arnaud Maurel. I thank seminar participants at the Ateliers de la DEPP, Séminaire interne du SIES, Rencontres d'Aussois and Labor and Public Economics seminar at PSE.

a 2.2 percentage points increase in vocational track students' probability of holding a long-term contract, and an increase of up to 4.5 percent in annual wages for some subgroups of vocational track students, namely male and high-achieving students.

Keywords: *vocational education, student tracking, returns to education, time to degree*

Introduction

Many countries are exerting efforts to increase the average schooling attainment level of their population, with the central objective of improving the labor market outcomes of low-skilled workers. In Europe, the strategic framework for European cooperation in education and training has set the objectives of reducing the share of early school leavers among the 18- to 24-year-olds, and increasing the share of 30- to 34-year-olds who have completed some form of higher education by 2020. These objectives have been integrated into a broader framework for reducing educational inequalities that aims to improve equality of opportunity and to tackle concerns about rising wage differentials. This context has fostered the implementation of educational policies targeted at vocational students, such as the opening of UCT hybrid schools in the UK in 2010 (Machin et al., 2020) or the 2008 reform of the vocational track in France. While they represent a large share of high school students, vocational track students are overrepresented among early school leavers and are overwhelmingly from low-income backgrounds.

Whether vocational students would be better off by attaining a higher level of qualification, however, remains a debated question. From a wage returns perspective, the answer to this question depends on the magnitude of the returns - as measured by wage differentials - of the considered educational investment compared to its cost. Exploiting changes in the institutional features of the educational system as a source of quasi-experimental variation to estimate the returns to schooling, the economic literature has provided compelling evidence of the large positive returns to education for early school leavers, with estimates of the causal impact of an additional year of schooling on wages falling in the 5–15 percent range.¹ Evidence on the labor market returns to education for the marginal enrollees beyond the compulsory schooling age is less clear-cut (see, for instance, Heckman et al. (2018)). Although a majority of studies concludes to positive returns for a wide range of marginal students, vocational students seem to stand apart from other marginal enrollees.² Leveraging reforms that increased the length of the vocational track in the 1950s to 1970s in Scandinavian and Eastern European countries, the literature tends to find limited effects of increased schooling on vocational students' labor market outcomes (see, for instance, Hall (2016); Zilic (2018); Bertrand et al. (2019); Oosterbeek and Webbink (2007)).

¹See, for instance, Angrist and Krueger (1991); Harmon and Walker (1995); Oreopoulos (2007); Meghir and Palme (2005); Duflo (2001); Aakvik et al. (2010) in secondary education.

²For evidence on the returns to education for the marginal enrollees to higher education, see, for instance, Maurin and McNally (2008); Zimmerman (2014); Carneiro et al. (2011).

Taking advantage of a reform of the French vocational track that increased vocational students' access to secondary education credentials and higher education, this paper provides new empirical evidence on the returns to schooling for this important subgroup of students.

The reform altered the structure of the high school graduation path for vocational track students by converting the former four-year programs into three-year programs, thereby reducing the time to high school graduation by one year and decreasing the opportunity cost of access to higher education. The reform was gradually rolled out over the period 2008 to 2011, generating within-middle-school variation in the degree of exposure to the converted programs over time. I exploit this variation to identify the impact of vocational programs conversions on vocational students' educational and labor market outcomes, and deal with selection by relying on a Bartik-type instrumental variable strategy that predicts exposure to the converted programs at the middle school level.

Using a unique set of administrative educational and labor market data, I estimate the impact of the reform on a wide range of educational and labor market outcomes. I find large positive effects of the reform on vocational track students' secondary schooling attainment, as measured by their probability of accessing Grade 12 (+14.8 percentage points) and their rate of *baccalauréat* completion (+10.0 percentage points). These effects on educational attainment were achieved despite an average 0.34 year decrease in time spent in secondary education. In addition, high school dropout and grade retention increased by 2.3 percentage points and 3.6 percentage points, respectively, but, unlike for other educational outcomes, these effects were only transitory and confined to the years of implementation of the reform.

The reform increased vocational students' access to higher education by 3.1 percentage points, representing a 28.4 percent increase from the baseline of 10.9 percent. Most of the reform's impact on access to higher education was achieved thanks to technical post-secondary programs (+2.2 percentage points for the first year and +1.4 percentage point for the second year of the programs, respectively). The reform had a positive impact on access to the first year of university programs (+1.7 percentage point), but almost no effect on the probability of access to the second year (+0.2 percentage point). The differential persistence of vocational track students across the two types of higher education programs translates into differential effects of the reform in terms of access to technical post-secondary vs. university credentials. While the reform had no effect on credits accumulation nor degree completion rates in university programs, it increased vocational track students' probability of graduating from post-secondary technical programs by 0.9 percentage

point.

The reduced-form effects of the reform on the labor market outcomes of vocational track students nine years after Grade 9 are positive but moderate. Their probability of holding a long-term labor market contract increased by 2.2 percentage points as a result of the reform, from a baseline of 50.7 percent. However, the reform did not have any effect on their average gross hourly wage, which remained of 12.2 euros, i.e., only slightly above the minimum hourly wage level (10.25 euros). The annual wage of vocational track students fully exposed to the reform increased by 1.8 percentage point, an effect that is not, however, statistically significant.

I assess the heterogeneity in the labor market effects of the reform based on students' unobservable and observable characteristics. Quantile regressions reveal no differential effects of the reform at different points of the distribution of annual wages. The effects for female vs. male students and for high vs. low achievers are more contrasted. Male students benefited from a large positive effect on their probability of holding a long-term contract (+3.3 percentage points), which is associated with a substantial increase in their annual wage levels (+5.4 percentage points), whereas the labor market outcomes of their female counterparts were essentially unaffected by the reform. I provide suggestive evidence that the heterogeneity in the labor market effects of the reform by gender can be explained by the lower premium attached to *baccalauréat* graduation in female-dominated vocational specializations compared to male-dominated specializations. The labor market outcomes of the 50 percent highest achievers among vocational track students were also significantly improved by the reform, with a 3.0 percentage points gain in their probability of holding a long-term contract and a 4.5 percentage points increase in their annual wages. On the contrary, no detectable effect can be observed for the 50 percent lowest achievers.

I investigate a range of potential channels that could have driven the educational effects of the reform, and provide suggestive evidence that the observed effects were mediated by the time to high school graduation channel alone. I further explore the heterogeneity in the treatment to which vocational track students were exposed by distinguishing between those with a high vs. a low predicted probability of graduating from high school in the absence of the reform. The former group of students was only exposed to a reduction in the time to graduation, while the latter was exposed both to a reduction in the time to graduation and to its associated effects on access to Grade 12 and *baccalauréat* completion. I perform a heterogeneity analysis based on this pre-reform student type, and find that the labor market effects of the reform were two to three times larger for

the former group of students. However, the loss of precision induced by the sample splits prevents me from detecting significant effects. Overall, these findings suggest that for the vocational track students who attain the high school graduation level, the returns to an additional year of secondary schooling are low enough to be more than offset by the associated loss of labor market experience.

Related Literature. Motivated by the finding that low-income students make different educational decisions compared to their high-income peers with the same achievement levels (Hoxby and Turner, 2015), a large strand of the literature has advocated for the need to consider how a number of financial, institutional, informational and behavioral factors can interfere with low-income students' decision to invest in education, both at the extensive and at the intensive margins. This vast body of research has shown that low-income students are at the same time less likely to be able to afford the costs of education (Angrist, 1993; Bettinger, 2004; Fack and Grenet, 2015; Dynarski et al., 2018), and more likely to face a range of adverse psychological mechanisms that affect their decision to invest in education, among which debt aversion (Field, 2009) or high time preferences (Oreopoulos, 2007). Insufficient access to information regarding their own achievement level (Guyon and Huillery, 2021), the returns to the different types of programs (Jensen, 2010; Hastings et al., 2015), the application processes (Castleman and Page, 2015) and the financial aid to which they are entitled (Bettinger et al., 2012) further contribute to low-income students' underinvestment in higher education. The behavioral economics literature has also emphasized the role that peer effects and social conformism play in shaping students' educational choices (Guyon and Huillery, 2016; Akerlof and Kranton, 2000; Dalton et al., 2016).

One important determinant of the decision to invest in education that has received less attention, however, is time to graduation. As the staged structure of the schooling system induces students to consider their educational pathway as a sequence of completed programs rather than as a sequence of independent years, time to graduation is likely to act as a major barrier to the educational attainment of low-income students. Indeed, the length of a program affects at the same time the direct costs of program completion (e.g., tuition fees or school materials cost) as well as its opportunity cost (the sum of the annual labor incomes foregone). To my knowledge, this paper is one of the first to investigate how time to graduation in secondary education affects students' schooling trajectories and entry-level labor market outcomes.³ One notable exception is Brodaty et

³For examples of analyses of time to graduation in higher education, see for instance Bound et al. (2012); Garibaldi et al. (2012).

al. (2008), who find negative returns to time to graduation, in a context where variation in time to graduation is entirely attributable to grade retention and may therefore act as a negative signal on the labor market. The context studied in this paper allows to answer a very different research question, that is the impact of institutional time to graduation, i.e., the theoretical time to high school graduation determined by the structure of secondary education.

This paper also relates to the literature on the returns to schooling. Little is known about the determinants of vocational track students' returns to schooling. Previous research has investigated the heterogeneity in the returns to general vs. vocational skills for vocational students, concluding to higher returns to vocational education or a hybrid form of education (Meer, 2007; Malamud and Pop-Eleches, 2010; Hampf and Woessmann, 2017; Mane, 1999; Machin et al., 2020). However, the comparative advantage of vocational education has been found to be decreasing with age (Hanushek et al., 2017). A recent study by Saltiel (2020) finds, based on a discrete choice model, that the higher returns to vocational education relative to general education in Switzerland can be explained by the strong linkage between vocational schooling and tertiary studies in this country. This paper is closely related to the present study, as it assesses the role played by higher education in determining vocational students' returns to schooling. The present study however differs from Saltiel (2020)'s, as it focuses on the French context and analyzes the returns to a mix of vocational and general education for vocational track students.

Other determinants of vocational students' returns to schooling have been much less studied. In particular, while program completion and certification – two dimensions that are directly related to time to graduation – have been shown to be major determinants of the returns to schooling (Oreopoulos and Petronijevic, 2013; Grenet, 2013; Ost et al., 2018), very few papers have assessed how they affect vocational students' returns to schooling. One notable exception is Alfonsi et al. (2020), which provide suggestive evidence on the matter by showing that the returns to certified school-based training are higher than the returns to uncertified firm-based training in the context of Uganda. By contrast, the present study gives evidence on the returns to the *baccalauréat* for school-based vocational track students.

The remainder of this paper is organized as follows. Section 1.1 describes the structure of the French schooling system and provides information on the institutional background of the reform and its implementation. Section 1.2 presents the empirical strategy. Section 1.3 describes the data sources used in this paper and provides summary statistics on vocational and non-vocational

students. Section 1.4 presents graphical evidence on the impact of the reform, before turning to the effects of exposure to the converted programs on vocational track students' secondary schooling attainment, access to higher education and credentials, and, ultimately, labor market outcomes. Section 1.5 discusses the mechanisms underlying the reform's effects while Section 1.6 concludes.

1.1 Institutional Background

In this section, I describe the structure of the French schooling system, before exposing the main features and implementation of the vocational track reform.

1.1.1 The French Schooling System

Student tracking in France. In France, education is compulsory between the ages of 6 and 16. After five years of primary school, students enroll in a common four-year middle school path.⁴ After completing middle school, students are tracked into different pathways depending on their achievement level and preferences. The general and technological track, on the one hand, consists of programs that lead to high school graduation within three years, through the completion of the technological or the general *baccalauréat*. Its main purpose is to prepare students for tertiary-level studies. The vocational track, on the other hand, consists of two-year vocational programs that lead to vocational graduation and three- to four-year programs leading to the vocational *baccalauréat*. The vocational track has the dual purpose of preparing students to higher education studies and to the school-to-work transition. Each year, approximately 25 percent of Grade 9 students enter the vocational track, while approximately half of the students join the general and technological track. The remaining 20 percent repeat Grade 9, join an apprenticeship program, or drop out of school.⁵

The vocational track. The vocational track consists of different types of vocational training programs, which differ in terms of their vocational specialization and of whether they are more or less intensive in vocational vs. general training. Among the more than 100 existing vocational specializations, one can distinguish between trades specializations (industrial engineering, agriculture and natural landscapes, industrial transformation, civil engineering and soft materials) and services

⁴A limited number of low-achieving students can be pre-tracked into specialized programs where they receive individual learning support starting in Grade 6 (SEGPA), or, at a latter stage, in a vocational type of Grade 9.

⁵Due to data limitations, apprentices cannot be distinguished from dropouts in this study. As a consequence, I adopt a definition of vocational track students that covers school-based vocational track students only.

specializations (management, communication, care and community services). The typical vocational track student receives 450 to 550 hours of vocational training and 250 to 360 hours of general training a year. The general training includes French, history and geography, civic and moral education, math, physics, foreign language, arts and sports courses. In addition to the school-based training, each student undertakes a mandatory six to seven-week internship once a year.

Academic national examinations. During their secondary schooling, students can take two academic national examinations. At the end of middle school, the vast majority of a cohort takes the *diplôme national du brevet* (DNB) national examination. In French and math, the DNB examinations are anonymously graded and therefore allow to assess the achievement level of the students on a comparable basis before high school entry. The students who attain Grade 12 are tested for the second time at the end of high school, with the *baccalauréat* examinations: the vocational, technological or general Grade 12 completion and success on the final *baccalauréat* examinations are certified by the vocational, technological or general *baccalauréat*, respectively.

The higher education system. Access to higher education is conditioned upon having obtained any of the three types of *baccalauréat*. First-year post-secondary programs include university programs, which are mostly non-selective, and selective programs.⁶ The latter fall into three main categories, with a strict academic stratification: two-year post-secondary vocational and technical programs (*sections de techniciens supérieurs* and *instituts universitaires de technologie*), post-secondary management and engineering schools, and the two-year elite *classes préparatoires aux grandes écoles* (CPGE), which prepare their students to the entrance examinations of the *grandes écoles*, the French elite graduate schools.

The vast majority of vocational students who attain higher education enroll in an STS or in a university program. The STS programs are selective two-year vocational and technical programs that are located in high schools. Each STS has a specific specialization and therefore trains its students for a limited range of vocational occupations. At the end of the two-year curriculum, STS students take the *brevet de technicien supérieur* (BTS) examination, and successful students are awarded the corresponding BTS degree. At university, vocational students generally enroll in three-year programs, called *licence*, which provide general training and prepare for the master's

⁶Since the 2018 reform of admission into higher education, universities are now allowed to select students according to their academic performance, but this study covers the pre-2018 period.

level rather than for a given occupation. Students are awarded academic credits (ECTS) for every university course that they successfully complete. When the student reaches 60 credits – the equivalent of 1500 to 1800 hours of work – she can pass to the next grade. At the end of the three years, the students having accumulated 180 credits are awarded the *licence* diploma.

1.1.2 The 2008 Vocational Track Reform

The high school vocational track before the reform. Panel (a) of Figure 1.1 describes the pre-reform structure of the different high school tracks in France. Before the reform, the vocational track consisted of two-year lower programs, the *Certificat d'aptitude professionnelle* (CAP) and the *Brevet d'études professionnelles* (BEP), and two-year upper programs, the vocational Grades 11 and 12. The BEP programs had a stronger general training component than the CAP programs, but both types of programs did not lead to the *baccalauréat* degree. To obtain the *baccalauréat* and hence be granted access to higher education, vocational students were required to undertake the two-year vocational Grades 11 and 12 programs after the completion of their BEP or, more rarely, of their CAP. The path to the *baccalauréat* was therefore a two-stage four-year curriculum, that I hereafter denote as the 2+2-year BEP path.⁷

Before the reform, more than 80 percent of vocational track students would enroll in a BEP, while the remaining 20 percent would enroll in a CAP (see Figure 1.2). Less than half of the BEP students would enroll in Grade 11, and only one third would complete the vocational *baccalauréat*. Among the CAP students, only 10 percent would enroll in Grade 11, and five percent would complete the vocational *baccalauréat*. Hence the vast majority of vocational track students would complete vocational education only.

Content of the 2008 reform. The 2008 reform modified the vocational high school graduation path as a means to reduce the time to high school graduation for vocational track students. A three-year integrated program, the *seconde professionnelle*, replaced the former two-year BEP programs, thereby reducing the time to high school graduation from four to three years for vocational track students. The possibility of being awarded a vocational certification after the completion of the two first years in the vocational track was maintained.⁸

⁷As only a marginal share of pre-reform CAP students would attempt to enroll in Grades 11 and 12, I hereafter do not distinguish between two-year CAP and 2+2-year CAP students, but rather refer to CAP students as a whole.

⁸The nature of this certification was, however, altered: while the pre-reform BEP certification was a *diplôme*, the post-reform BEP certification became a *certification*, and therefore no longer has any legal value on the labor market.

Similarly to other large scale institutional reforms, the 2008 vocational track reform came with a number of side adjustments. The supply of vocational specializations was adapted to the local demand for skills-specific labor while the average class size of the converted vocational programs (*seconde professionnelle*) was reduced. In Section 1.5, I show that these specific features of the reform are unlikely to have mediated the reform's effects on vocational students' educational and labor market outcomes, and that the treatment received by vocational track students can be interpreted as resulting from the reduction in the time to high school graduation.

Timing of the reform. The reform was gradually implemented between the years 2008 and 2011.⁹ This implementation was monitored at the education region (*académie*) level, but high schools were to a large extent free to decide when to convert their BEP into *seconde professionnelle* programs.¹⁰ This generated important variation in exposure to the converted programs across middle schools, districts (hereafter *départements*) and *académies* in a given year, and within each of these geographical units over time. Of the 7,500 BEP programs that existed in 2007, 26 percent were converted into *seconde professionnelle* programs in 2008, 55 percent in 2009, 4 percent in 2010 and 14 percent in 2011 (see Figure 1.3).

A number of features of the two-year BEP influenced the timing of conversions across high schools. Table 1.1 reports the odds ratios for the probability that a BEP was converted in 2009, 2010 or 2011 compared to 2008, for a range of BEP baseline characteristics, including their socio-demographic composition, their pre-reform rates of access to Grade 11 and to the *baccalauréat*, and their vocational specialization. The programs that were converted earlier tended to be programs whose rates of access to Grade 11 or to the *baccalauréat* were higher, and, on average, had better-performing students. The table also reveals clear conversion patterns according to the vocational specialization of the BEP, with earlier conversions among the service-oriented BEP programs than among the industry-oriented. The more machinery involved in the training, the more challenging it was for the school principals to reorganize a four-year schedule into three years. This explains why the civil engineering specializations were 3.6 times more likely than other specializations to be converted in 2009 compared to 2008, and the industrial specializations were 1.6 times more likely to be converted in 2009 compared to 2008, compared to other specializations. One specialization stood

⁹The new *seconde professionnelle* programs were experimented at a very small scale between 2005 and 2007, with roughly one percent of vocational track students enrolling in an experimental *seconde professionnelle* in these years.

¹⁰In France, the *académie* is a regional educational district that is responsible for the implementation of the national schooling policies.

apart from other services specializations: the care specializations were systematically converted in 2011, because the corresponding *baccalauréat* degree did not exist before the reform.

In 2008, and, to a lesser extent, in 2009, the coexistence of the former two-year BEP and the converted programs gave substantial room to the selection of students into the different types of programs. To alter their odds of being admitted to a two-year BEP vs. a *seconde professionnelle*, prospective vocational students may have switched their choice of specialization or, more exceptionnally, might have adopted more drastic strategies to escape the reform, for instance by applying to CAP programs only. These adjustments are unlikely to have occurred in a random manner. One can think, for instance, that among the students willing to graduate from high school, those facing more severe budget constraints might have been more inclined to enroll in a *seconde professionnelle* program and therefore save one year of educational costs compared to the 2+2-year BEP option. By contrast, students with a poor achievement level may have felt unsure about their ability to complete the *baccalauréat* within three years, and might therefore have opted for the safer option of enrolling in a 2+2-year BEP.

In the next section, I expose the empirical strategy, which deals with potential adjustments in students' high school application strategies by making use of a prediction of their level of exposure to the reform at the middle school level.

1.2 Empirical Strategy

In this section, I present the empirical strategy and discuss its identifying assumptions.

1.2.1 An IV Approach

To identify the impact of the vocational track reform on students' outcomes, I exploit the staggered adoption of the reform that generated differential exposure to the converted programs across middle schools in a given year, and within middle schools over time, and deal with selection into the converted programs by relying on a Bartik-like IV approach.

The reduced-form equation of interest is as follows:

$$Y_{i,m,t} = \alpha + \beta * \widehat{Exp\acute{o}sure}_{m,t} + \theta_m + \theta_t + \epsilon_{i,m,t} \quad (1.1)$$

where $Y_{i,m,t}$ is the outcome of student i from middle school m in year t , $\widehat{Exp\acute{o}sure}_{m,t}$ is the

predicted exposure to converted programs of students from middle school m in year t , θ_m and θ_t are middle school and year fixed effects, respectively, and $\epsilon_{i,m,t}$ is an error term that is assumed to be uncorrelated to $Exp\hat{posure}_{m,t}$ conditional on the other regressors. As explained below, $Exp\hat{posure}_{m,t}$ is a weighted average of the share of converted programs in the vocational high schools serving a student's middle school m in year t . The parameter β therefore identifies the effect of switching from 0 to 100 percent the share of converted programs in the vocational high schools serving a student's middle school, under the "common trend" assumption. When this assumption is verified, i.e., when the underlying trends in students' outcomes are not systematically correlated with a middle school's level of exposure to the reform, the variation in the outcomes of students from middle schools whose degree of exposure to the converted programs remains stable after the reform can be used as a counterfactual for the evolution for the outcomes of students from middle schools whose exposure to the converted programs increased after the reform. The comparison of each middle school's observed vs. counterfactual evolutions then allows to retrieve the reform's effect. The standard errors are clustered at the year \times middle school level in order to allow for arbitrary correlation between the unobserved determinants of the outcomes of students from the same middle school in the same year.

As discussed in the previous section, actual enrollment in *seconde professionnelle* programs is likely to be endogenous to the students' observable and unobservable characteristics, if e.g. more able students felt more confident to enroll in the converted programs, while less able students were more likely to enroll in a CAP instead. To overcome this issue, I predict each student's level of exposure to the converted programs based on the pre-reform enrollment patterns of students from the same middle school after Grade 9. Let $\gamma_{h,t}$ denote the share of converted programs in vocational high school h in year t and $\omega_{m,h,t}$ the share of middle school m 's students who enroll in high school h in year t . For each middle school m and post-reform year t , the predicted exposure to the reform is defined as follows:

$$Exp\hat{posure}_{m,t} = \sum_h \omega_{m,h,t_0} \gamma_{h,t},$$

where t_0 denotes the last pre-reform year (i.e., 2007).

Given the many potential channels of the reform discussed in Section 1.1, I do not compute the two-stage least-squares (2SLS) coefficients obtained by instrumenting matriculation in *seconde professionnelle* with exposure to the *seconde professionnelle* programs, but rather stick to the reduced-form approach. Note that the strength of the first stage, which is of almost 90 percent

(see Section 1.5), implies that the reduced form and associated 2SLS coefficients are of similar magnitude. In addition, Section 1.5 provides evidence that the effects of the reform were mostly mediated by the time to high school graduation channel.

1.2.2 Identifying Assumptions

Common trend assumption. The common trend assumption might be violated for a number of reasons. The reform required considerable adjustments from the part of the high schools' staff. Teachers had to learn how to prepare the students for *baccalauréat* completion in three years instead of the previous four. In particular, they had to find the proper balance between moving at a faster pace to make sure that most students would be adequately prepared for the *baccalauréat* examinations, and limiting the risk of dropout induced by a too intensive teaching for the weakest students.¹¹

Several features of the high schools are likely to have influenced the timing of adoption of the reform, among which the motivation, the flexibility and the quality of the high school staff. Middle schools with a positive pre-trend in access to Grade 12 and *baccalauréat* might have been more exposed to high schools whose staff was more eager to opt for an early conversion of the BEP programs. For instance, the presence of a dynamic high school headmaster might have both contributed to drive up the average outcomes of a middle school in the pre-reform years and at the same time have facilitated the adoption of the reform. Such violation of the common trend assumption can be indirectly tested by controlling for middle school specific trends.

A more serious threat to the identification strategy would be the existence of dynamic treatment effects, which would lead to a violation of the common trend assumption in the years of implementation of the reform. For instance, assume that the teachers learned from their teaching experience with their first *seconde professionnelle* students, and could better prepare the subsequent cohorts of students. In such a case, treated students from later cohorts are likely to have obtained better outcomes, on average, than treated students from earlier cohorts. As this phenomenon would be observed in the absence of variation in exposure to the reform, the OLS estimation of the model described by Equation (1) would incorrectly attribute the corresponding variation in outcomes to

¹¹This adjustment is likely to have been particularly complex, and resulted in a learning process throughout the years of implementation of the reform for the teachers in the vocational track. Consistent with this, Figure 1.B1 in Appendix indeed documents a transitory increase in high school dropout in the vocational track during the years of implementation of the reform.

the “spontaneous” evolution of outcomes for treated students, leading to an underestimation of the reform’s impact. In Section 1.4.5, I show that the results are robust to the use of an alternative estimator, the Wald-TC estimator proposed by de Chaisemartin and d’Haultfoeuille (2019). This estimator relies on a weaker version of the common trend assumption that allows for dynamic treatment effects.

Validity of the Bartik instrument. $Exp\hat{posure}_{m,t}$ is the product of a vector of “shocks” received by high schools, $\gamma_{h,t}$, and of a vector of “shares” measuring the weight that each middle school assigns to each high school, $\omega_{m,h,t}$. $Exp\hat{posure}_{m,t}$ therefore belongs to the family of Bartik instruments. The validity of the Bartik instrument can be assessed in terms of the exogeneity of the shares, the shocks, or both.

Exogeneity of the shares. As shown by Goldsmith-Pinkham et al. (2020), traditional asymptotics that rely on the number of high schools and time periods being fixed while the number of middle schools goes to infinity imply that the validity of the Bartik instrument can be interpreted as solely based on the exogeneity of the shares $\omega_{m,h,t}$. As argued by the authors, this assumption is more likely to hold if all shares are set to their pre-reform values, i.e. if the Bartik instrument uses ω_{m,h,t_0} for all years. This implies that only pre-reform students’ tracking decisions are used to predict post-reform students’ exposure to the converted programs, which is the approach used in this study.

Table 1.2 presents summary statistics regarding the shares that are used to compute the Bartik instrument. Although entirely similar to a Bartik instrument from an accounting point of view, the instrument used in this paper exploits a design that deviates slightly from the traditional framework where shocks (growth rates) are received by national industries, and used to estimate the inverse elasticity of labor supply. In the present study, the shocks (the rate of BEP conversions) are received by high schools, that are not akin to industries, due to their stronger local component. By design, ω_{m,h,t_0} , the share of middle school m ’s pre-reform students enrolling in high school h is equal to 0 for the vast majority of middle schools m . Unlike in the traditional setting that exploits shocks to industries, it therefore appears very unlikely that a small number of high schools could be driving the estimates.¹²

¹²In their paper, Goldsmith-Pinkham et al. (2020) show that the Bartik instrument can be decomposed into a weighted sum of just-identified instrumental variables estimators, with weights measuring the sensitivity of the parameter to each instrument. Therefore, reporting information on the instruments (the shares) receiving the highest weights allows to provide transparent information on the variation in the data that drives the estimates. For computational reasons, I limit my argumentation to the reasons that make me confident that my estimates are not driven by a limited number of schools, without reporting the formal tests suggested by Goldsmith-Pinkham et al.

Exogeneity of the shocks. Alternatively, the validity of the Bartik instrument can be assessed in light of the exogeneity of the shocks. Measurement error could reintroduce endogeneity, as the shocks can only be observed when the *seconde professionnelle* class was actually opened, which happened when demand for this class was sufficient. Most likely, when demand for a new *seconde professionnelle* program was insufficient, the corresponding BEP was closed rather than maintained for an additional year.¹³ This solution cannot have been massively adopted, as the number of BEP/*seconde professionnelle* programs did not incur any significant drop in the reform years (see Figure 1.3). It may, however, have prevailed locally, which could bias some middle schools' estimated treatment effects. To overcome these concerns, I check in Section 1.4.5 that the results are robust to using the predicted exposure to the converted programs at the *département* and *académie* levels instead of at the middle school level.

1.3 Data and Summary Statistics

This section presents the data sources used in this paper, and provides some summary statistics on vocational vs. general and technological track students.

1.3.1 Data Sources

To evaluate the impact of the vocational track reform, I combine several exhaustive administrative micro datasets that provide detailed information on students' educational trajectories in secondary and higher education as well as their labor market outcomes at an early stage of their career. For the first time with French data, I was able to match the educational and labor market administrative databases using a nominative statistical matching procedure that uniquely matches individuals based on their names, first names, and date and place of birth. The details of the statistical matching procedure are reported in Appendix 1.A (see in particular Table 1.A1 for some statistics on the quality of the statistical matching).

(2020). The Stata package proposed by the authors to recover the weights requires data in wide format with $h \times t$ columns for the shares, the shocks, and the product of the two. With $h = 1977$ and $t = 3$, this implies an almost 18,000 columns dataset, which is not computationally feasible.

¹³As the demand for a given *seconde professionnelle* program can only be observed *after* the submission of students' choices, it appears unlikely that high schools had room to eventually change the program type but still manage to fill the class outside of the official assignment procedure.

Enrollment data. Students' enrollment status in secondary education for the years 2006 to 2017 is measured from the SYSCA administrative datasets of the French Ministry of Education. Along with basic sociodemographic characteristics including gender, social background and date and place of birth, these data contain information on the students' school, Grade level, and program of matriculation. The SYSCA data also covers the students enrolled in selective higher education programs that are located in high schools, namely the STS and CPGE programs. I supplement these data with the SISE administrative data from the French Ministry of Higher Education, which cover the students enrolled in most of the other higher education programs (see Appendix 1.A for more details).

Secondary education examination data. I enrich the dataset with information on the students' examination results in secondary education, using the OCEAN databases managed by the French Ministry of Education.

I use the OCEAN-DNB database to measure students' ability prior to entering the vocational track. This database reports the grades obtained in each of the subjects that are tested with the national DNB examination. For math and French, the written examinations are blind tests, i.e. anonymously and externally graded, and students' performance in these subjects can be considered as an objective measure of their relative achievement level at the end of middle school. I compute the students' percentile ranks in French and math separately by year of examination and *académie*, to account for the fact that the examination files are specific to each *académie*.

The OCEAN-BAC database reports, for all candidates to the *baccalauréat* examinations, the type and series of *baccalauréat* along with the grades obtained in each subject. I use this database to assess how the vocational track reform impacted vocational high school graduation, as measured by the completion of the vocational *baccalauréat*.

Higher education degrees. I use the OCEAN-BTS database from the French ministry of Education to determine whether STS students obtained the BTS degree. For university programs, I retrieve from the SISE databases information on the ECTS credits accumulated by students and on *licence* graduation.

Labor market outcomes. Eventually, I enrich the dataset with information on vocational students' labor market outcomes between 2016 and 2018 using the 2017 and 2018 Insee DADS

(*Déclaration annuelle de données sociales*) databases. These yearly administrative databases cover the universe of public and private wage earners working or living in France in the considered and the preceding year. For every wage earner, the data contain information on the different periods of employment, the number of hours worked and the gross wages earned. For each period of employment, a qualitative description is also available, with information on the type of contract and the socioeconomic category of the job.

Outcomes of interest. All educational outcomes of interest are computed within a fixed number of years after Grade 9, for them to be comparable across cohorts. Secondary education outcomes are measured within four years after Grade 9. I build dummy variables indicating whether the student accessed Grade 12, completed the vocational *baccalauréat* and whether she dropped out from high school.¹⁴ Higher education outcomes are measured within five to seven years after Grade 9. I distinguish between the access to STS programs and to university programs, which are the two main types of post-secondary programs attended by vocational students.

For labor market outcomes, I rely on three different point measurements, since only two DADS files, covering three years of labor market data, are available. I measure students' gross annual wage, gross hourly wage as well as the type of labor contract nine years after Grade 9, for the Grade 9 cohorts 2007 to 2009.

1.3.2 Sample Definition and Summary Statistics

Sample definition. The sample consists of all students who attended Grade 9 in France between 2006 and 2014. I hereafter refer to these students as the 2006 to 2014 cohorts, where cohort n refers to the students who attended Grade 9 for the first time in the schooling year $n - 1/n$, and whose theoretical year of entry into high school is the schooling year $n/n + 1$.¹⁵ Out of the 6,576,000 students in the considered cohorts, 1,608,072 joined the vocational high school track after Grade 9.

In the main analysis, I restrict the sample to the 2006 to 2009 cohorts, in order to assess the reform's impact on outcomes up to nine years after Grade 9 on a balanced sample. This restricted sample consists of 2,841,198 students, of whom 688,742 joined the vocational track after Grade 9.

¹⁴To ensure the time consistency of the definition of high school dropout, I define high school dropout as the non-completion of the first two years of a vocational program.

¹⁵Considering all Grade 9 students rather than vocational track students allows to evaluate the potential effects of the reform on students' tracking decisions after Grade 9 (e.g. strategic Grade 9 repetition, or the decision to join the vocational vs. the general and technological high school track).

Summary statistics. Columns 1 and 2 of Table 1.3 contrast the characteristics of vocational and non-vocational track students, for the cohorts 2006 to 2009. Vocational track students account for roughly one quarter of a cohort and are typically males, are disproportionately from low to medium-low SES and are more likely to be low achievers.

There are, however, substantial disparities among vocational track students, in particular across the various types of programs offered in the vocational track. Columns 3, 4 and 5 of Table 1.3 compare vocational track students across the three different types of programs that existed before the reform: the CAP, the 2-year BEP and the 2+2-year BEP. Before the reform, CAP, 2-year BEP and 2+2-year BEP students accounted for respectively 17, 43 and 40 percent of vocational track students, with a clear stratification by ability between the different types of programs, as indicated both by the share of students who did not take the DNB examination (a strong signal of low achievement level at the end of middle school) and the average percentile rank of DNB candidates. To this ability stratification corresponds a social stratification, with, on average, more high-SES students in 2+2-year BEP than in 2-year BEP, and in 2-year BEP than in CAP—all types of programs however receiving on average more than 50 percent of low-SES students. Female students are slightly more numerous in CAP and 2+2-year BEP than in 2-year BEP.

1.4 Results

In this section, I provide graphical evidence on the impact of the reform, before turning to the IV approach.

1.4.1 Graphical Evidence on the Impact of the Reform

Figures 1.4 to 1.6 plot the cohort averages of the main outcomes of interest for the 2006 to 2014 cohorts of vocational track and non-vocational track students separately, while Table 1.4 compares the 2007 and 2009 cohorts. The analysis reveals that the reform led to a very large increase in the average secondary schooling attainment of vocational track students and to a more moderate increase in their average higher education attainment and relative improvement in their labor market outcomes nine years after Grade 9, over a period during which the outcomes of non-vocational students were relatively stable.

Outcomes in secondary education. Figure 1.4 highlights contrasted dynamics in the evolution of educational outcomes for vocational vs. general track students in the pre-reform period, the 2007–2011 period, and the post-reform period. Before 2007 and after 2011, the secondary education outcomes of vocational track and non-vocational track students followed a parallel and slightly positive trend. In contrast, the 2007–2011 period was marked by a large improvement in vocational students’ educational outcomes relative to non-vocational students. Their rate of access to Grade 12 increased from 36 percent to 62 percent, representing a 72 percent increase compared to the baseline, and their probability of completing a vocational *baccalauréat* increased from 27 to 42 percent, representing a 56 percent increase from the baseline. During the same period, non-vocational students’ probabilities of accessing Grade 12 or completing the *baccalauréat* increased only smoothly, with increases of 9 percent and 11 percent from their baseline values, respectively. These patterns rule out the conjunction of other large scale policies or circumstances that would have impacted all students during this period. In addition, the fact that the differential evolution in vocational and non-vocational outcomes is confined to the years of implementation of the reform further suggests that the observed evolutions in vocational track students’ outcomes can be safely attributed to the vocational track reform.

Outcomes in higher education. The reform’s impact on vocational students’ higher education outcomes is less spectacular in absolute value, but corroborates the previous findings: the trend in vocational track students’ outcomes deviated from that of other students in the years of implementation of the vocational track reform (see Figure 1.5). Between 2007 and 2011, their probability of accessing the first year of an STS increased from 8 percent to 12 percent (a four percentage points increase, representing a 50 percent increase from the baseline) while non-vocational students’ access rate remained stable. The same phenomenon can be observed for the access to the second year of an STS, with a two percentage points increase for vocational students. Vocational students’ access to university does not appear, however, to have been substantially affected by the reform.

Outcomes on the labor market. At first sight, the figures displayed in Table 1.4 suggest that nine years after Grade 9, vocational students’ labor market insertion is comparable to that of non-vocational students, with an employment rate close to 70 percent. This apparent similarity, however, hides heterogeneous situations: nine years after Grade 9, 12 percent of non-vocational

students are still enrolled in education, which is the case of only a marginal share of vocational track students (0.8 percent). To facilitate the comparison, individuals still enrolled in education are excluded from the computation of labor market outcomes.

Before the reform, 50.9 percent of vocational track students would hold a long-term contract nine years after Grade 9, against only 46.1 percent of non-vocational students. The corresponding gap between vocational and non-vocational students is higher in 2009 (6.6 percentage points) than in 2007 (4.8 percentage points).

Overall, this descriptive analysis suggests that the vocational track reform led to large improvements in vocational students' secondary education outcomes, and more limited in magnitude (but substantial compared to their baseline values) effects on their access rate to higher education and persistence in post-secondary programs. Additionally, it appears plausible that the reform had small positive effects on the labor market outcomes of vocational track students. However, the limited number of years available for the analysis of labor market outcomes prevents from analyzing the extent to which the observed variations can be attributed to the reform. To isolate the sole effect of exposure to the converted programs, I now turn to the instrumental variable approach.

1.4.2 Validity of the IV Strategy

In this section, I argue that the sample of analysis can be safely restricted to vocational students when estimating the impact of exposure to the converted programs on students' outcomes. I then provide graphical evidence supporting the validity of the proposed instrumental variable strategy, by comparing the evolutions over time in the outcomes of vocational students with a high versus low probability of being exposed to early BEP conversions.

Sample definition. Before the reform, close to a quarter of Grade 9 students would enroll into the vocational track in the following year, of which 80 percent would matriculate in a BEP. As a consequence, in the absence of reform-induced changes in the matriculation decisions of the students, the expected take-up of converted programs for Grade 9 students would be of at most 20 percent. To avoid this low take-up setting, I restrict the sample to vocational students and show that this restriction is unlikely to induce sample selection bias, as the reform does not appear to have affected the share of students joining the vocational track after Grade 9, as evidenced by the tests reported in Section 1.A.3 of Appendix 1.A. In addition, I show in Section 1.4.5 that the

results are robust to extending the sample to the entire Grade 9 cohort.

Wage earners subsample. Since the wage data cover only the years 2016 to 2018, I can compare the labor market outcomes n years after Grade 9 for only three consecutive cohorts. I therefore define a subsample that consists of the wage earners of the cohorts 2007 to 2009, for whom the data allow to measure labor market outcomes nine years after Grade 9. The validity of this restriction relies on the absence of compositional effects of the reform on vocational track students that are wage earners nine years after Grade 9. I formally test this assumption in Section 1.A.3 and find no evidence of a reform-induced selection effect into the wage earners databases.

IV approach: graphical evidence. Although the proposed IV strategy exploits the whole available variation in predicted exposure to the reform, I adopt in this section a simplified setting in which I compare the average outcomes of middle schools exposed early vs. late to the vocational track reform.¹⁶ This simple comparison provides clear graphical evidence on the role that BEP conversions played in determining the observed evolution of vocational track students' outcomes (see Figures 1.7 to 1.9). Panel (a) of Figure 1.7 shows that a middle school's predicted exposure to early conversions is a strong predictor of its Grade 9 students' probability to enroll in a *seconde professionnelle*. In 2008, 45 percent of students from the early-exposed middle schools matriculated in a *seconde professionnelle* the following year, compared to only 5 percent among students from the middle schools that were later exposed. Between 2008 and 2009, the reverse trend can be observed, as the second group of middle schools became exposed to BEP conversions. The enrollment rates in *seconde professionnelle* stabilize for students from both groups of middle schools between 2009 and 2010, reflecting the absence of BEP conversions in this period. Eventually, the final round of BEP conversions that took place in 2011 triggered similar increases in the probabilities of matriculation in *seconde professionnelle* for the two groups of students in 2011.

The evolutions of the secondary and higher education outcomes for the two groups of students largely reflect this differential timing of exposure to the converted programs. While outside of the reform years, the outcomes of both groups follow very similar trends, they diverge between 2007 and 2009. In 2008, the trend followed by early exposed students deviates from that of later exposed students for all the outcomes of interest, with important increases in their rates of access

¹⁶I consider as early exposed the middle schools belonging to the top quartile of predicted exposure to the reform in 2008, and later exposed those belonging to the bottom quartile of predicted exposure in 2008.

to Grade 12, *baccalauréat* completion, and probability of enrolling in the first or second year of a post-secondary program. In 2009, the outcomes of later exposed students fully (for secondary education outcomes) or partially (for higher education outcomes) catch up that of the early exposed.

These findings lend strong support to the validity of the identification strategy. First, pre- and post-reform trends in the evolutions of outcomes of middle schools that were exposed to the reform at different points in time are very similar. As a consequence, using later-exposed middle schools as a counterfactual for early-exposed middle schools appears to be a sound strategy. Second, the predicted exposure to the reform is a strong predictor of matriculation in a *seconde professionnelle* program. Therefore, predicted exposure to the BEP conversions can be considered as a powerful instrument for observed matriculation in *seconde professionnelle*.

1.4.3 Main Results

Educational outcomes. Table 1.5 and Figure 1.10 report the reduced-form coefficients on predicted exposure to the reform obtained from the estimation of Equation (1.1) on the 2006–2009 cohorts of vocational students. The coefficients measure the effects of switching from 0 to 100 percent the share of converted programs in the vocational high schools serving a Grade 9 student’s middle school, either using the baseline specification (Column 1) or controlling for individual characteristics, middle school specific trends, and the two sets of controls at the same time (Columns 2 to 4). Since the coefficients are very robust to the inclusion of each of these controls, I hereafter comment on the coefficients from Column 1 only.

Outcomes in secondary education . The reform had an unambiguously large positive impact on the secondary schooling attainment of vocational track students: their likelihood of attaining Grade 12 increased by 14.8 percentage points from a baseline of 35.5 percent (+42 percent), associated to an increase in the probability of completing a vocational *baccalauréat* of 10.0 percentage points, from a baseline of 26.5 percent (+38 percent), both coefficients being statistically significant at the one percent level. These large effects on educational attainment were achieved despite an average 0.34 year decrease in time spent in secondary education. High school dropout and grade retention rates increased by respectively 2.3 percentage points and 3.6 percentage points, but, unlike for all previously listed outcomes, these effects were only transitory and confined to the years of implementation of the reform (see Appendix 1.B for a discussion of the methodological issues related to the time-consistency of the definition of these two outcomes, and Figure 1.B1 for the

overall evolutions of these two outcomes during the 2006 to 2014 period).

Outcomes in higher education. The reform significantly increased vocational students' probability of accessing higher education and of accumulating post-secondary vocational credits. However, it had no impact on their likelihood of completing a university degree, nor on their average accumulation of ECTS credits in university programs. Vocational students' probability of accessing higher education increased by 3.1 percentage points from a baseline of 10.9 percent (+28 percent), of which two thirds correspond to an increased probability of accessing an STS program (2.2 percentage points, from a baseline of 7.8 percent), and one third to an increased probability of accessing a university program (1.7 percentage point, from a baseline of 2.6 percent), both coefficients being significant at the one percent level. Vocational students' probability of accessing the second year of a post-secondary program increased by 1.3 percentage point—this effect being driven by the increased probability of accessing the second year of an STS program. The reform also increased vocational students' likelihood of obtaining an STS degree by 0.9 percentage point, from a baseline of 3.9 percent (+23 percent), but had very limited effects on degree completion and accumulation of ECTS credits in university programs: neither the probability of graduating from university nor the credits accumulation were significantly affected.

Labor market outcomes. Table 1.6 reports the reduced-form coefficients on predicted exposure to vocational programs conversions obtained from the estimation of Equation (1.1) on the 2007–2009 subsample of wage earners.

The reform had positive but moderate effects on vocational students' labor market outcomes nine years after Grade 9. Their probability of holding a long-term contract increased by 2.2 percentage points, from a baseline of 50.7 percent. This effect is associated with an insignificant 1.8-percent increase in their annual wages. As their hourly wages were not affected by the reform, these findings overall suggest that the moderate annual wage effects reflect a slight improvement in job stability (as suggested by the effects on long-term contracts), which translates into a higher number of hours worked during the year.

1.4.4 Heterogeneity Analysis

In this section, I analyze the heterogeneity in the reform's impact across groups of students with different observable and unobservable characteristics. Tables 1.C1 to 1.C3 in Appendix 1.C

and Figures 1.11 to 1.13 present the impact of the reform by gender, SES, and prior academic achievement level, whereas Figure 1.14 shows the results from the quantile regression of log annual wages on exposure to vocational program conversions.

Effects by gender and socioeconomic background. Male vs. female students' and high to medium-low SES vs. low-SES students' educational outcomes were similarly impacted by the reform.¹⁷ The reform's impact on labor market outcomes appears to be more differentiated, but the low precision of the estimates does not allow to detect statistically significant differences between the considered groups of students, at the five percent level.

Male students exhibit a positive effect on their probability of holding a long-term contract (+3.3 percentage points, significant at the one percent level) as well as on their annual wages (+5.4 percent, significant at the five percent level), while the corresponding effects for female vocational students are close to zero and not statistically significant.

Low-SES students' probability of holding a long-term contract increased by 3.5 percentage points (significant at the five percent level), this effect being associated with an insignificant 1.7-percent increase in their annual wages. By contrast, I find no evidence of significant labor market effects for medium-low- to high-SES students.

Effects by prior achievement level. The effects of the reform by prior achievement level are more contrasted.¹⁸ While the rates of access to Grade 12 of both above- and below-median achievers were similarly impacted by the reform, the reform had heterogeneous effects on their probabilities to complete a vocational *baccalauréat* or access higher education, as well as on their annual wages.

For above-median students, I find that full exposure to the reform induced a large increase in *baccalauréat* completion and access to higher education, by 9.9 and 4.1 percentage points, respectively, significant at the one percent level, and find positive effects on both their probability of holding a long-term contract (+3.0 percentage points, significant at the one percent level) and on their annual wage (+4.5 percent, significant at the five percent level).

By contrast, the results indicate only moderate effects on the educational outcomes of below-

¹⁷When assessing the impact of the reform by student socioeconomic background, I contrast the outcomes of low SES students, who account for roughly 50 percent of vocational track students, with those of medium-low to high SES students.

¹⁸I define two equal sized groups of students to compare the effects of the reform by prior achievement level, by grouping together the students who did not take the DNB examinations and the students who belong to the bottom quartile of the DNB percentile rank distribution, whom I compare to students in the second to fourth quartiles.

median achievers, whether we consider their probability of completing a vocational *baccalauréat* (+3.4 percentage points, significant at the one percent level), or their probability of accessing higher education. While below-median achievers were as likely to access university as those above the median, the increase in their probability of accessing an STS is twice lower (+1.5 percentage points vs. + 3.1 percentage points). These modest educational effects translate into an insignificant drop of 5.6 percent in their annual wages, and no significant effect on their probability of holding a long-term contract.

Heterogeneity on unobservables. The previous analyses have focused on the mean effect of exposure to the reform on annual wages. To uncover potential heterogeneous effects of the reform at different points of the wage distribution, I adopt a quantile regression framework. Let $Q_\tau(Y|X) = F^{-1}(\tau|X)$ denote the conditional quantile function, where $F_{Y|X}$ is the cumulative density function of Y conditional on X , and the vector X includes our measure of predicted exposure to the reform, $\widehat{Exposure}_{m,t}$, middle school and year fixed effects, and potential additional controls. The conditional quantile function is assumed to be a linear function of X , i.e.,

$$Q_\tau(Y|X) = X'\beta_\tau, \tag{1.2}$$

or equivalently

$$Y = X'\beta_\tau + \epsilon_\tau, \text{ with } Q_\tau(\epsilon_\tau|X) = 0, \tag{1.3}$$

Then,

$$\beta_\tau = \arg \min_{\beta} E[\rho_\tau(Y - X'\beta)], \tag{1.4}$$

where $\rho_\tau(u) = (\tau - \mathbb{1}\{u < 0\})u$.

Figure 1.14 displays the quantile regression estimates of the coefficients obtained on exposure to the reform for the first to ninth deciles of the wage distribution, controlling for middle school and year fixed effects. Although the figure shows a declining pattern of wage effects of the reform by decile of the wage distribution, the figure reveals no differential effects of exposure to the reform at the different deciles of the wage distribution. For all deciles, exposure to the reform induces a positive but insignificant increase in annual wages nine years after Grade 9.

This section has investigated the heterogeneity of the reform’s impact across different socio-demographic groups of students and at different points of the wage distribution. The results indicate that the main dimensions of heterogeneity in the labor market effects of the reform are the students’ gender and prior achievement level. Before digging into the potential mechanisms underlying the effects of the reform in Section 1.5, I assess the robustness of the results to a set of alternative specifications in the next section.

1.4.5 Robustness Checks

In this section, I assess the robustness of the results to relaxing the common trend assumption, extending the sample of analysis to non-vocational students, and using only the variation in exposure to the reform at the *département* or *académie* levels, rather than at the middle school level.

Alternative estimator. As discussed in Section 1.2, Equation (1.1) may yield biased estimates if the common trend assumption is violated during the years of implementation of the reform. To address this issue, I investigate the robustness of the results to relaxing the common trend assumption by using an estimator that allows for dynamic treatment effects. Specifically, I rely on the Wald-TC estimator proposed by de Chaisemartin and d’Haultfoeuille (2019), which compares the evolution between $t - 1$ and t of the outcomes of students from middle schools which had the same degree of exposure to the reform in $t - 1$. This group-wise comparison allows to control for the potential lagged effects of the treatment received in $t - 1$ (see Appendix 1.D for more details on the Wald-TC approach and on the parameter tuning implemented in the analysis). The Wald-TC estimates obtained on the wage earners subsample are reported in Table 1.D1. Reassuringly, they are very similar to those from the baseline specification; however, all coefficients on labor market outcomes become insignificant.

Extending the sample to non-vocational students. Table 1.D2 in the Appendix presents the results obtained when we extend the analysis sample to the entire cohorts of 2007 to 2009 Grade 9 students. This allows to control for any potential impact of the reform on the tracking decision in or out of the vocational track (which I have shown in section 1.2 to be very limited in any case). The main cost of this alternative approach is that the “first stage” is weaker. However, all reduced-form coefficients of interest are of the same order of magnitude as those obtained with

the main specification.

Predicting exposure to the reform at the *département* and *académie* levels. Tables 1.D3 and 1.D4 in the Appendix present the results obtained when estimating Equation (1.1) using predicted exposure to the converted programs at the local area (*département*) and at the schooling authority (*académie*) level, rather than at the middle school level. Again, the results obtained for each outcome of interest are very similar to those from the main specification.

1.5 Discussion

In this section, I first investigate a range of potential mechanisms that could have mediated the reform's impact, and provide evidence that the main driver of the reform's effects was the reduction in time to graduation (Section 1.5.1). Second, I acknowledge the heterogeneity in the nature of the time to graduation treatment across the different types of pre-reform students, and implement an empirical strategy to disentangle the direct effect of the reform on the length of schooling from its indirect effects on secondary schooling attainment (Section 1.5.2). Lastly, I provide suggestive evidence that the differential returns to the *baccalauréat* across male-dominated vs. female-dominated vocational specializations can at least partly explain the heterogeneity of the reform's impact by gender (Section 1.5.3).

1.5.1 Potential Mechanisms

To better understand the mechanisms through which the reform could have impacted vocational students' educational and labor market outcomes, I investigate a number of potential channels by estimating the effects of the reform on vocational track matriculation, vocational specialization decisions, and first year vocational programs' class size, using the baseline empirical specification (Equation (1.1)). The results, which are reported in Appendix Table 1.E1, suggest that matriculation in *seconde professionnelle* programs can be considered as the main channel explaining the reform's impact on students' outcomes.

Matriculation in vocational track. As expected, exposure to the reform largely influenced students' decision of whether or not to enroll in a *seconde professionnelle* program, with an 85 percentage points increase in the probability to do so for fully exposed students compared to

unexposed students. Matriculation in *seconde professionnelle* programs cannot however *a priori* be considered as the only mediator of the reform's effects. When highly exposed to the converted programs, students were also on average induced to matriculate more often in CAP programs (+3.1 percentage points).

To assess the importance of the CAP channel in explaining the overall impact of the reform, I estimate Equation (1.1) on a subsample that excludes the *départements* with the highest contribution to the CAP channel.¹⁹ Table 1.E2 in the Appendix presents the corresponding coefficients, obtained on approximately 90 percent of the observations in the initial sample. While the CAP channel can by construction be ruled out for this subsample, the estimated effects of the reform on all of the outcomes of interest are highly similar to the main specification's findings.

Class size. The reform significantly reduced class size in the first year of vocational programs, with an average reduction of almost 2 students per class, from a baseline average class size of 13.7 students. I show, however, that class size is unlikely to be an important channel for the observed effects of the reform on students' outcomes. I use the same approach as for the investigation of the CAP matriculation channel, i.e., I estimate the impact of the reform on a subsample of *départements* whose exposure to the converted programs is not associated with class size reductions, and report the coefficients in Table 1.E3 in the Appendix. As expected, no significant effect of the reform on class size can be detected for this subsample, which consists of 70 percent of the observations in the initial sample, but all coefficients on the outcomes of interest are similar to the main specification's results, though slightly (but insignificantly) lower for most educational outcomes and higher for labor market outcomes. These findings are consistent with Leuven et al. (2008) and Piketty et al. (2006) who find limited effects of class size reduction in secondary education on school performance, as well as with Leuven and Løkken (2020) and Falch et al. (2017), who find no significant effects on wages.

Vocational specialization. I find very limited effects of increased exposure to the reform on the distribution of vocational students across the different vocational specializations. I only find a 1.3 percentage points decrease in the probability of opting for a civil engineering specialization, which is significant at the five percent level, but is not robust to the inclusion of controls.

¹⁹These *départements* are identified by performing leave-one-out regressions of Equation (1.1) for the matriculation in CAP outcome. The *départements* whose omission contributes to a significant reduction of the CAP matriculation coefficient are then removed from the sample.

1.5.2 Impact of the Reform by Pre-Reform Student Type

While the introduction of the *seconde professionnelle* programs had the same impact on the theoretical time to high school graduation for all vocational students, it had heterogeneous implications on the schooling attainment level for the different pre-reform types of students. For the 2+2-year BEP students who would complete the *baccalauréat*, the reform consisted in a reduction of the time to graduation. By contrast, the reform affected both time to graduation and the secondary schooling attainment for the pre-reform two-year BEP students and the 2+2-year BEP students who would have failed to complete the program absent the reform. In this section, I investigate whether this treatment heterogeneity is associated with different labor market effects of the reform.

Predicting students' pre-reform type. Before the reform, vocational track students could enroll in a CAP, in a two-year BEP, or in a 2+2-year BEP. I predict each student's pre-reform type, i.e., the type of vocational program she would have undertaken absent the reform, based on a penalized multinomial logistic regression, using as explanatory variables the students' middle school of origin, the type of Grade 9 program attended (Grade 9, Grade 9 SEGPA, etc...), age, gender, parental occupation, and prior achievement level. Details of the estimation procedure are provided in Appendix 1.E.

Heterogeneous treatment effects by students' pre-reform type. Table 1.E4 presents the impact of the reform separately for each predicted student type. As expected, the reform had very limited effects on predicted CAP students, since CAP programs were only indirectly and marginally impacted by the reform. I therefore limit the comparison to the predicted 2-year BEP vs. predicted 2+2-year BEP students (defined as those who completed the program), for which I find very different effects of the reform.

Predicted 2-year BEP students' secondary education outcomes were largely improved by vocational programs conversions, with a 17.3 percentage points increase in their access rate to Grade 12 and a 9.9 percentage points increase in their rate of *baccalauréat* completion. Though high in magnitude, these results nevertheless indicate that almost half of the students who were induced by the reform to reach Grade 12 failed to complete the *baccalauréat*. These educational effects are associated with a positive effect on the probability of holding a long-term contract (+1.7 percentage point, significant at the five percent level) and a positive but insignificant impact on annual wages

(+1.2 percent).

To the contrary, for predicted 2+2-year BEP students, the effects on the probability to attain Grade 12 or complete the *baccalauréat* are of the same order of magnitude (9.5 percentage points and 8.4 percentage points, respectively).²⁰ The point estimates indicate that these students' probability of holding a long-term contract increased by 3.7 percentage points and that their annual wages increased by 4.2 percent. These effects, however, are not statistically significant at all conventional levels.

Overall, these results suggest that the positive impact of the reform on vocational students' labor market outcomes was most pronounced for students who would have obtained the *baccalauréat* in the absence of reform and who gained one year of labor market experience thanks to the reduction in time to graduation. It is important to note, however, that the students who would not have obtained the *baccalauréat* under the four-year curriculum also gained from the reduction in time to graduation. Since the reform is found to have improved these students' secondary schooling attainment, these results suggest that their returns to secondary schooling nine years after Grade 9 were high enough to offset the loss of one year of labor market experience.

In the next section, I investigate whether the returns to female-dominated vs. male-dominated specializations can explain the observed heterogeneity in the labor market effects of the reform by gender.

1.5.3 Exploring the Gender Heterogeneity in the Labor Market Effects of the Reform

As the degree of gender segregation across vocational specializations is extremely high in the vocational track, heterogeneous returns to the *baccalauréat* across vocational specializations could explain the finding that the reform improved the labor market outcomes of male vocational students but not of females (see Section 1.4.4). To investigate this possibility, this section compares the employment probabilities, annual wages, and hourly wages of pre-reform *baccalauréat* graduates and non-graduates in female-dominated vs. male-dominated specializations. Although these observed differences do not allow to infer on the causal returns to the *baccalauréat*, they provide suggestive

²⁰If the prediction accuracy was of 100 percent, one would expect to find no effect on the secondary schooling attainment of predicted 2+2-year BEP, as they are defined as the students who completed the curriculum. However, as can be noted in Column 5, only 59.7 percent of the predicted 2+2-year BEP students would complete the *baccalauréat* before the reform.

evidence that the limited labor market effects of the reform on female students is related to the smaller returns to the *baccalauréat* in female-dominated specializations.

Female vs. male-dominated vocational specializations. Table 1.E5 in the Appendix highlights the degree of gender segregation across vocational specializations. Out of nine broad vocational specializations, only one - management - has a balanced gender composition (45 percent of male students). With 72 to 97 percent of male students, industrial engineering, civil engineering, industrial transformation and agriculture and natural landscapes specializations, which all fall among trades specializations, are largely male-dominated. By contrast, community services specializations, care, communication and soft materials specializations, which are mostly services specializations, have between 81 and 94 percent of female students, and are therefore female-dominated.

Table 1.E6 shows the employment rates and the average annual and hourly wages of vocational students nine years after Grade 9, separately for female and male-dominated specializations. Female-dominated specializations exhibit at the same time lower employment rates, lower annual wages and lower hourly wages. This heterogeneity could be explained by the unequal returns to male vs. female-dominated specializations, but also by the self-selection of individuals with different levels of labor market aptitudes into male-dominated vs. female-dominated specializations, or by the differential preferences in terms of career-family trade-offs. Heterogeneity in labor market aptitudes is unlikely to explain the observed pattern. Indeed, on average, female-dominated specializations gather higher achieving students, as evidenced by the higher share of DNB candidates among the students and the higher average DNB percentile rank among DNB candidates (see Columns 2 and 3 of Table 1.E5). Therefore, for heterogeneity in unobserved labor market aptitudes to explain the lower wages of female-dominated specializations, these unobservables would need to be negatively correlated with ability, which would contradict both economics theory and available empirical evidence. As regards the impact of differential career-family balance choices, while heterogeneous preferences and constraints (e.g. parental leaves) could explain the lower annual wages of female-dominated specializations, they fail to explain their lower hourly wages. As a consequence, the wage gap between male and female-dominated specializations is likely to reflect, to a non-negligible extent, unequal returns to male-dominated vs. female-dominated specializations.

Comparison of *baccalauréat* graduates and non-graduates between female-dominated and male-dominated specializations. The comparison of *baccalauréat* graduates and non-

graduates between male-dominated vs. female-dominated specializations can help to understand the differential effect of the reform on male vs. female students. For all specializations, *baccalauréat* graduates exhibit higher employment probabilities, annual wages and hourly wages than non-graduates.²¹ However, in female-dominated specializations, the annual wages, and, to a larger extent, the hourly wages of *baccalauréat* graduates are much closer to those of non-graduates, compared to male-dominated specializations. This is particularly salient for the hourly wages, which are almost the same for *baccalauréat* graduates and non-graduates across all female-dominated specializations but one, while they are four to nine percent higher for *baccalauréat* graduates in male-dominated specializations.

Although purely descriptive, this evidence suggests that the *baccalauréat* wage premium is higher for male-dominated compared to female-dominated specializations. These findings are consistent with Saltiel (2020) who finds higher returns to tertiary education for vocational students who opt for a trades-based training in secondary education (i.e., in the French context, male-dominated specializations), compared to those who opt for a services-based training (i.e., female-dominated specializations). As such, they can contribute to explain why the reform improved male students' labor market outcomes without altering those of female students.

1.6 Conclusion

Exploiting the staggered adoption of the 2008 reform of the French vocational high school track, I show that the conversion of the 2+2-year BEP programs into three-year *seconde professionnelle* programs resulted in a large improvement in vocational track students' average secondary education outcomes, as measured by their access rate to Grade 12 and their rate of *baccalauréat* completion. The reform also increased vocational students' access to higher education with, however, a clear divide between technical and university post-secondary programs in terms of student persistence and degree completion.

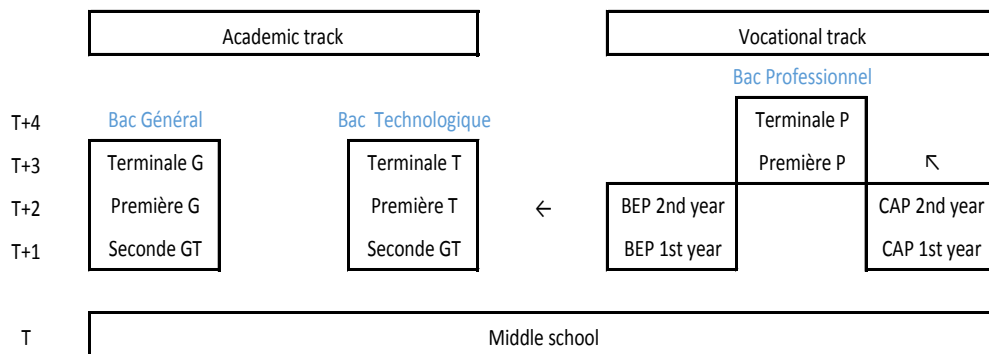
The reform had positive but moderate labor market effects. Vocational track students' job security, as measured by their probability of holding a long-term labor market contract nine years after Grade 9, improved, generating a small positive increase in their average annual wages, which is significant only for some subgroups of students, namely males and the top half achievers of the

²¹The natural landscapes and agriculture specializations stand apart, but they account for only a marginal share of pre-reform students.

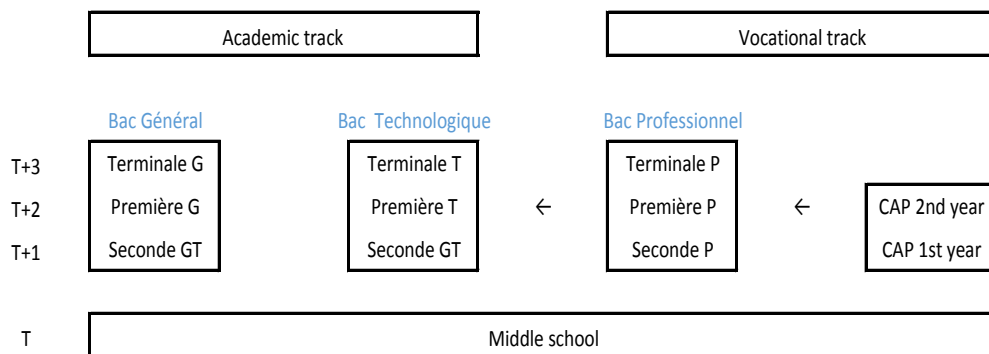
vocational track.

I provide suggestive evidence that the effects of the 2008 reform were entirely mediated by the introduction of the *seconde professionnelle* programs, and can therefore be interpreted as the effects of reducing from four to three years the time to high school graduation—a determinant of the schooling investment decision that has rarely been studied. By reducing the time spent at school by 0.34 year on average, the reform likely impacted vocational students' labor market experience. The analyses carried out in this study suggest that the reform-induced variation in labor market experience can partly explain the positive effects of the reform on the probability of holding a long-term labor market contract. However, the results also reveal heterogeneous labor market effects of the reform across groups of students who incurred similar reductions in their number of years of secondary schooling. I interpret this finding as evidence that the effects of the reform on the probability of access to the *baccalauréat* and to higher education are key to understanding its positive labor market effects.

Figure 1.1 – Structure of the French Secondary Education System



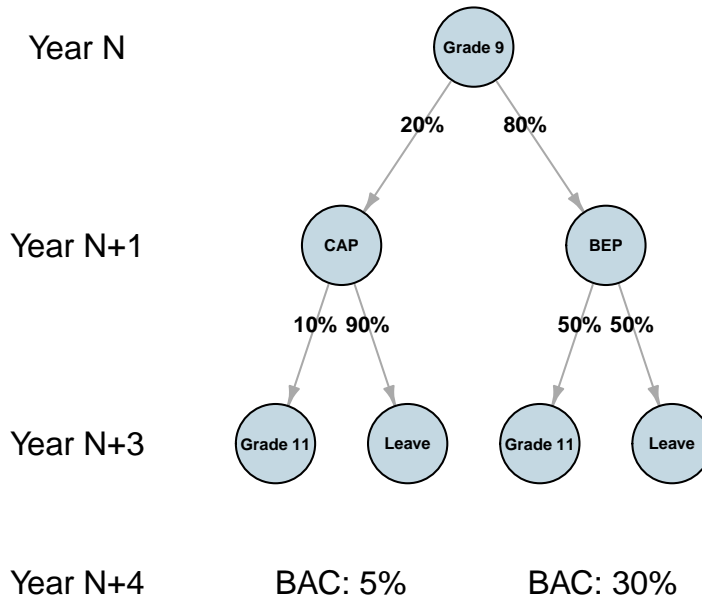
(a) Before the reform



(b) After the reform

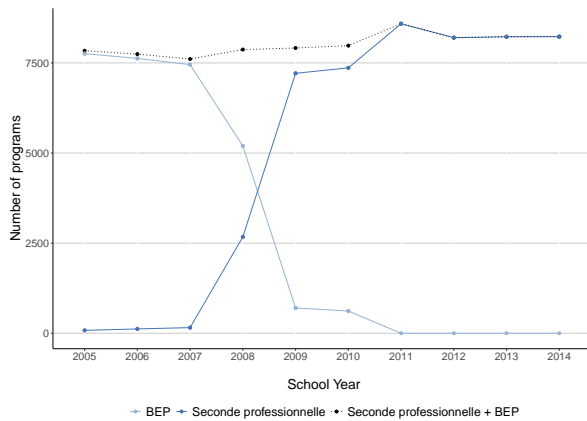
Notes: The diagram shows the structure of the high school tracks in France, before and after the 2008 reform of the vocational track. The academic track corresponds to the general and technological (GT) tracks. Arrows illustrate exceptional pathways between the different types of programs.

Figure 1.2 – Vocational Students’ Educational Pathways Before the Reform

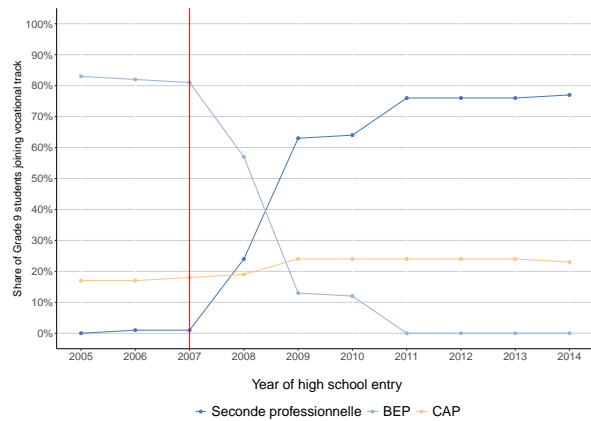


Notes: The diagram shows the share of pre-reform vocational track students attending and completing each considered pre-reform vocational track program. The figures are computed on the universe of students who were enrolled in Grade 9 between 2006 and 2007 and who attended a vocational track program the following year, based on administrative data from the French Ministry of Education.

Figure 1.3 – BEP Conversions and Vocational Students’ Enrollment Status After Grade 9



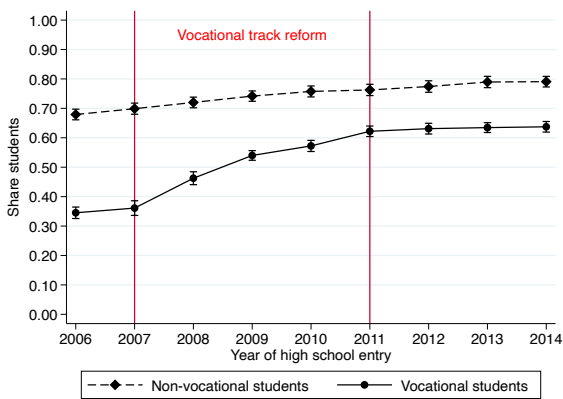
(a) Number of BEP and *Seconde professionnelle*



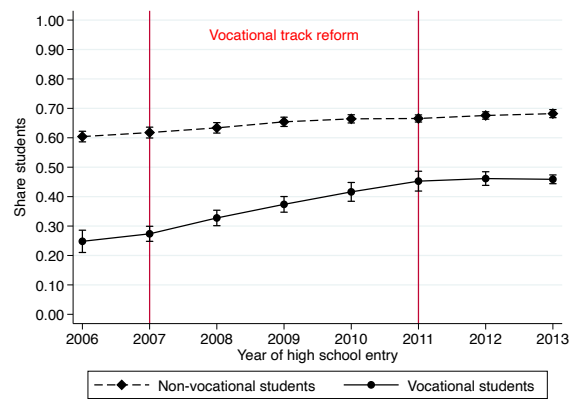
(b) Vocational students: enrollment status after Grade 9

Notes: Panel (a) displays the number of first-year BEP and *seconde professionnelle* programs operating each year in the vocational track. Panel (b) plots the share of vocational track students matriculating in a CAP, BEP, or *seconde professionnelle*, by year of high school entry. The figures are computed based on 2006 to 2014 administrative data from the French Ministry of Education.

Figure 1.4 – Graphical Evidence on the Impact of the Reform on Student Outcomes in Secondary Education: Vocational vs. Non-Vocational Students



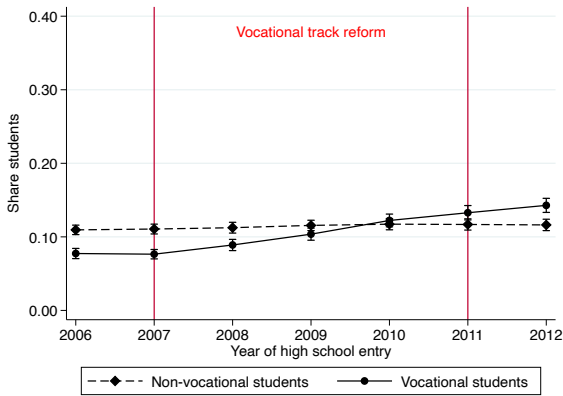
(a) Access to 12th grade



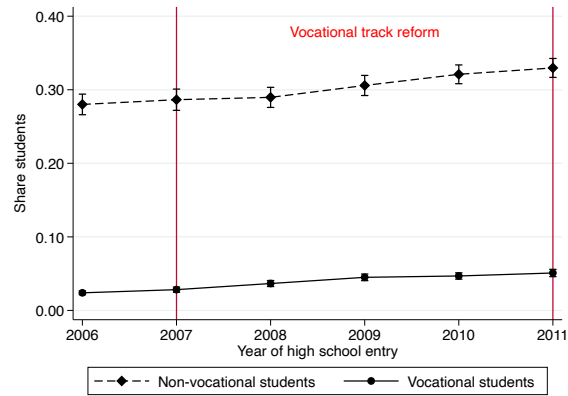
(b) Vocational *baccalauréat* completion

Notes: The figure displays the shares of students attaining the considered educational attainment level for vocational and non-vocational students separately, along with their 95 percent confidence intervals, by year of high school entry. The figures are computed on the universe of students who were enrolled in Grade 9 between 2006 and 2014, based on administrative data from the French Ministries of Education and Higher Education. The vertical red lines denote the last pre-reform year and the first post-reform year, respectively.

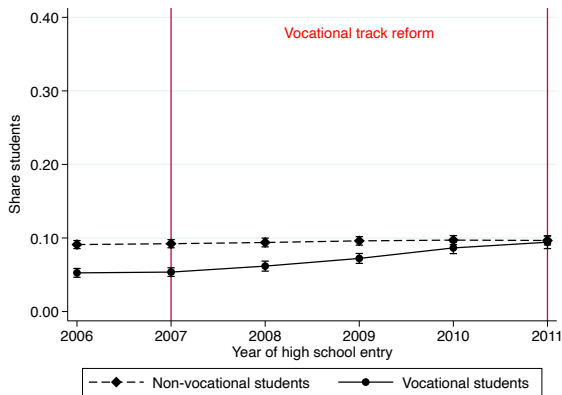
Figure 1.5 – Graphical Evidence on the Impact of the Reform on Student Outcomes in Higher Education: Vocational vs. Non-Vocational Students



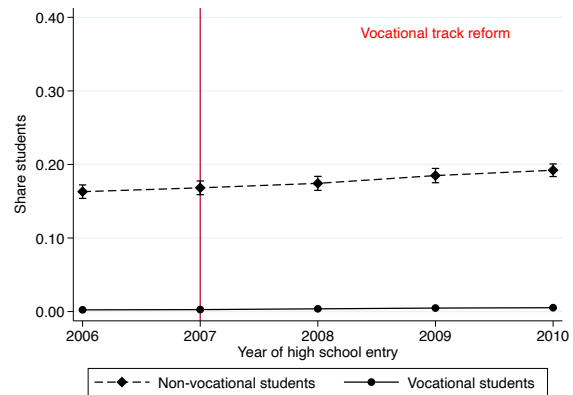
(a) Access to first year of STS



(b) Access to first year of university



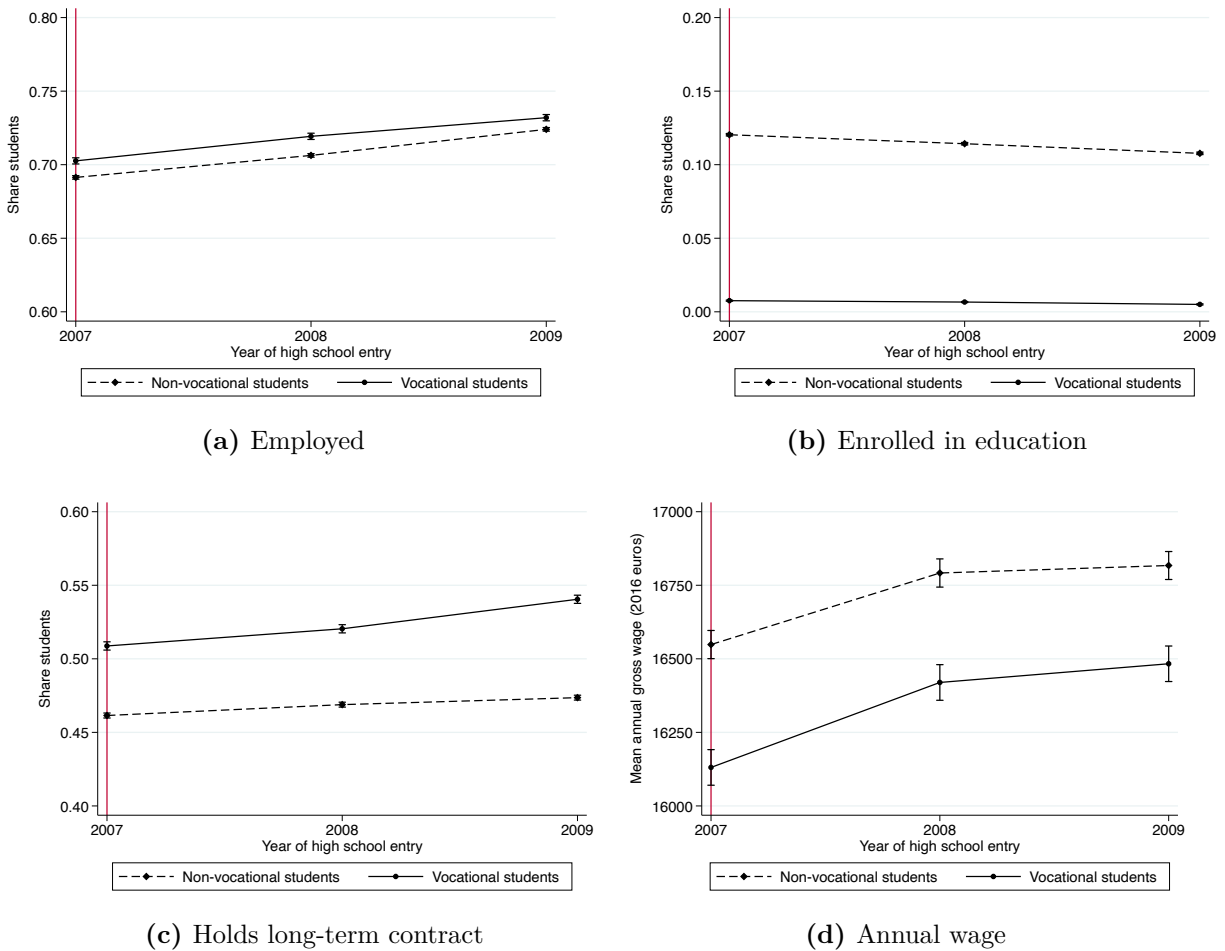
(c) Access to second year of STS



(d) Access to second year of university

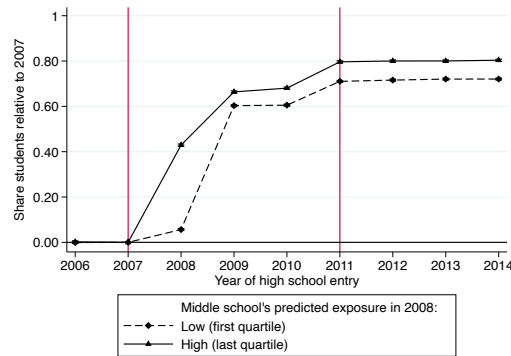
Notes: The figure displays the shares of students attaining the considered educational attainment level for vocational and non-vocational students separately, along with their 95 percent confidence intervals, by year of high school entry. The figures are computed on the universe of students who were enrolled in Grade 9 between 2006 and 2014, based on administrative data from the French Ministries of Education and Higher Education. The vertical red lines denote the last pre-reform year and the first post-reform year, respectively.

Figure 1.6 – Graphical Evidence on the Impact of the Reform on Labor Market Outcomes Nine Years after Grade 9: Vocational vs. Non-Vocational Students

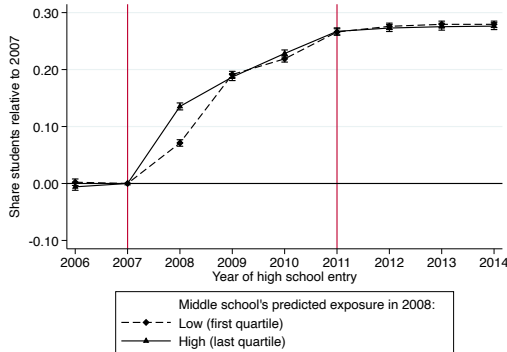


Notes: The figure displays the shares of students employed, enrolled in education, and holding a long-term labor market contract nine years after Grade 9, as well as the mean annual gross wage for vocational and non-vocational students separately, along with their 95 percent confidence intervals, by year of high school entry. The values from Panels (c) and (d) are computed on the subset of individuals not enrolled in education nine years after Grade 9. The figures are computed on the universe of students who were enrolled in Grade 9 between 2007 and 2009, based on administrative data from the French Ministries of Education and Higher Education and Insee.

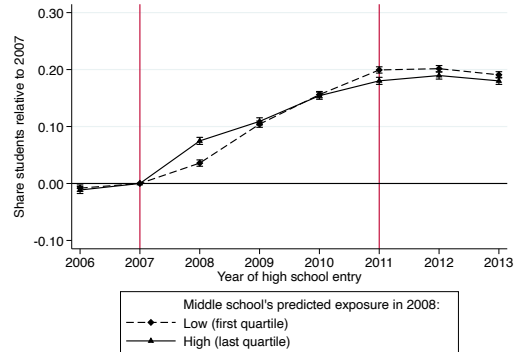
Figure 1.7 – Illustration of Empirical Strategy: Secondary Education Outcomes by Predicted Exposure to Early Vocational Program Conversions



(a) Enrolment in *seconde professionnelle*



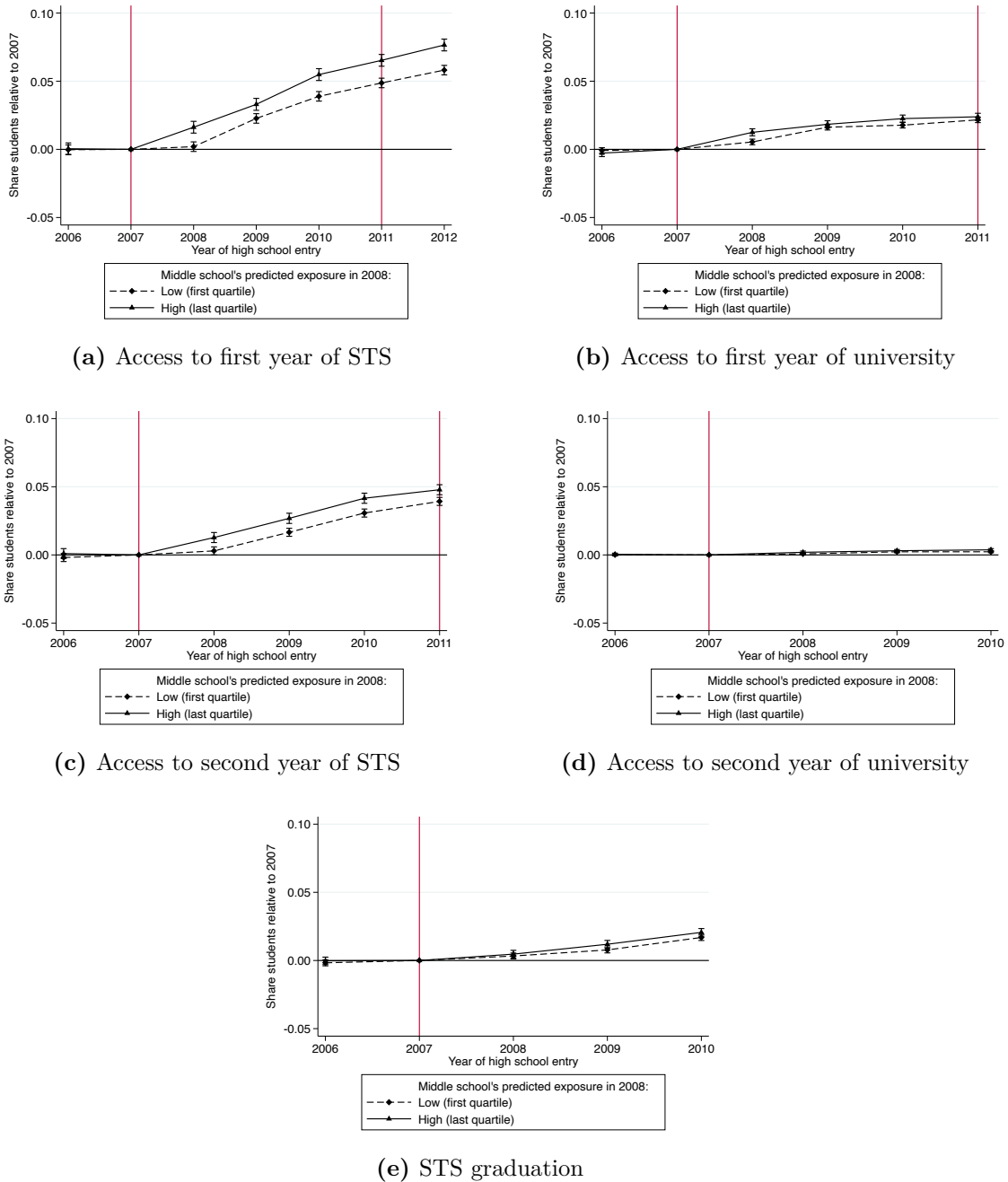
(b) Access to 12th grade



(c) Vocational *baccalauréat* completion

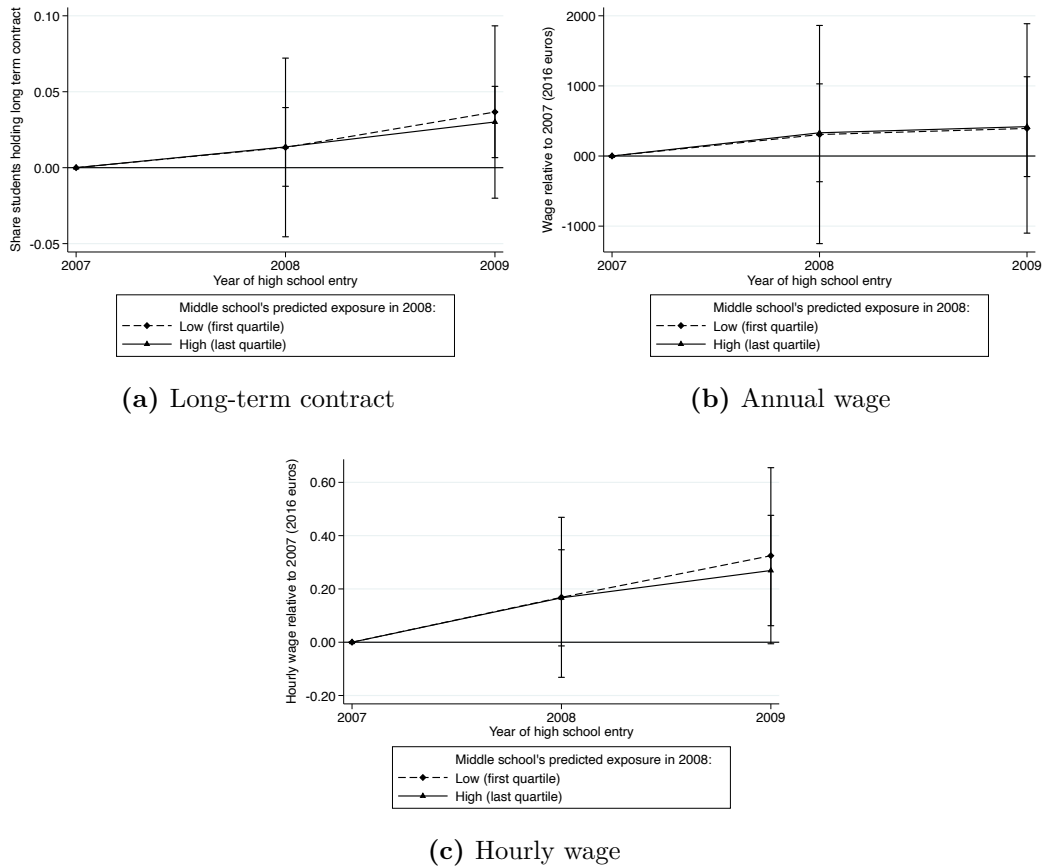
Notes: The figure displays the shares of vocational track students attaining the considered educational attainment level for students in middle schools that were exposed early vs. late to the vocational track reform, by year of high school entry. Early (later) exposed middle schools are defined as those belonging to the top (bottom) quartile of predicted exposure to BEP program conversions in 2008. The figures are computed on the universe of students who were enrolled in Grade 9 between 2006 and 2014 and who attended a vocational track program the following year, based on administrative data from the French Ministries of Education and Higher Education. The vertical red lines denote the last pre-reform year and the first post-reform year, respectively.

Figure 1.8 – Illustration of Empirical Strategy: Higher Education Outcomes by Predicted Exposure to Early Vocational Program Conversions



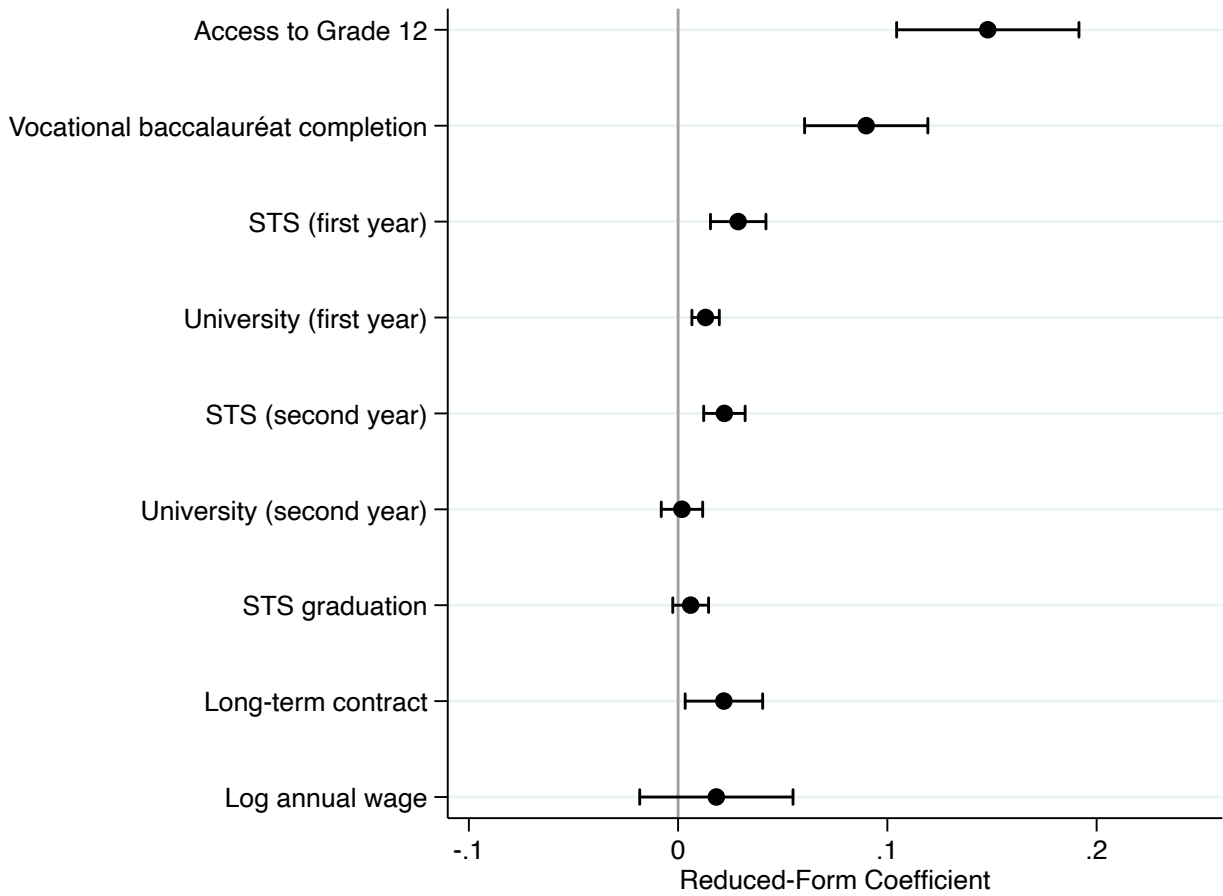
Notes: The figure displays the shares of vocational track students attaining the considered educational attainment level for students in middle schools that were exposed early vs. late to the vocational track reform, by year of high school entry. Early (later) exposed middle schools are defined as those belonging to the top (bottom) quartile of predicted exposure to BEP program conversions in 2008. The figures are computed on the universe of students who were enrolled in Grade 9 between 2006 and 2014 and who attended a vocational track program the following year, based on administrative data from the French Ministries of Education and Higher Education. The vertical red lines denote the last pre-reform year and the first post-reform year, respectively.

Figure 1.9 – Illustration of Empirical Strategy: Labor Market Outcomes Nine Years after Grade 9 by Predicted Exposure to Early Vocational Program Conversions



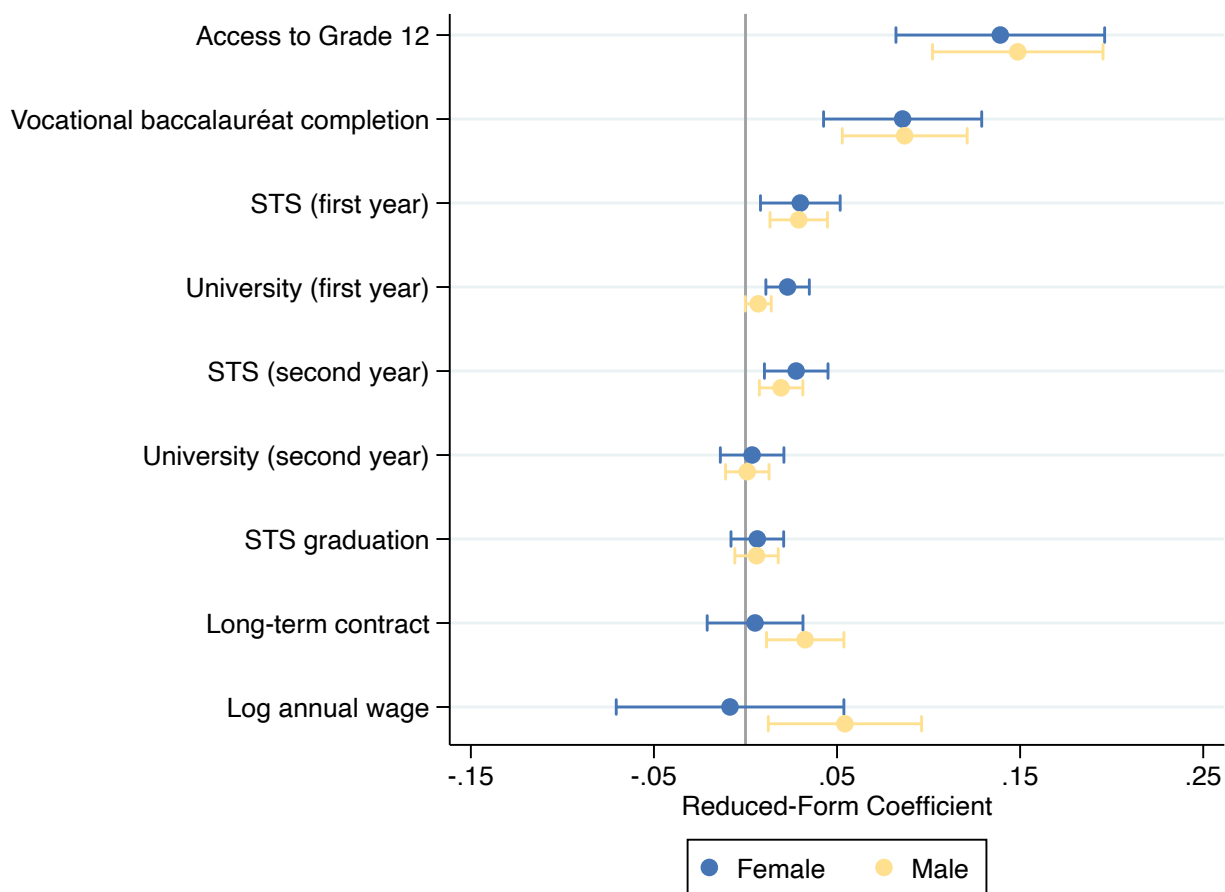
Notes: The figure displays the shares of vocational track students holding a long-term contract, as well as the average annual and hourly wages, for students in middle schools that were exposed early vs. late to the vocational track reform, by year of high school entry. Early (later) exposed middle schools are defined as those belonging to the top (bottom) quartile of predicted exposure to BEP program conversions in 2008. The figures are computed on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year and were wage earners in the DADS data 2016–2018, based on administrative data from the French Ministries of Education and Higher Education and Insee.

Figure 1.10 – Reduced-Form Coefficients on Exposure to Vocational Program Conversions



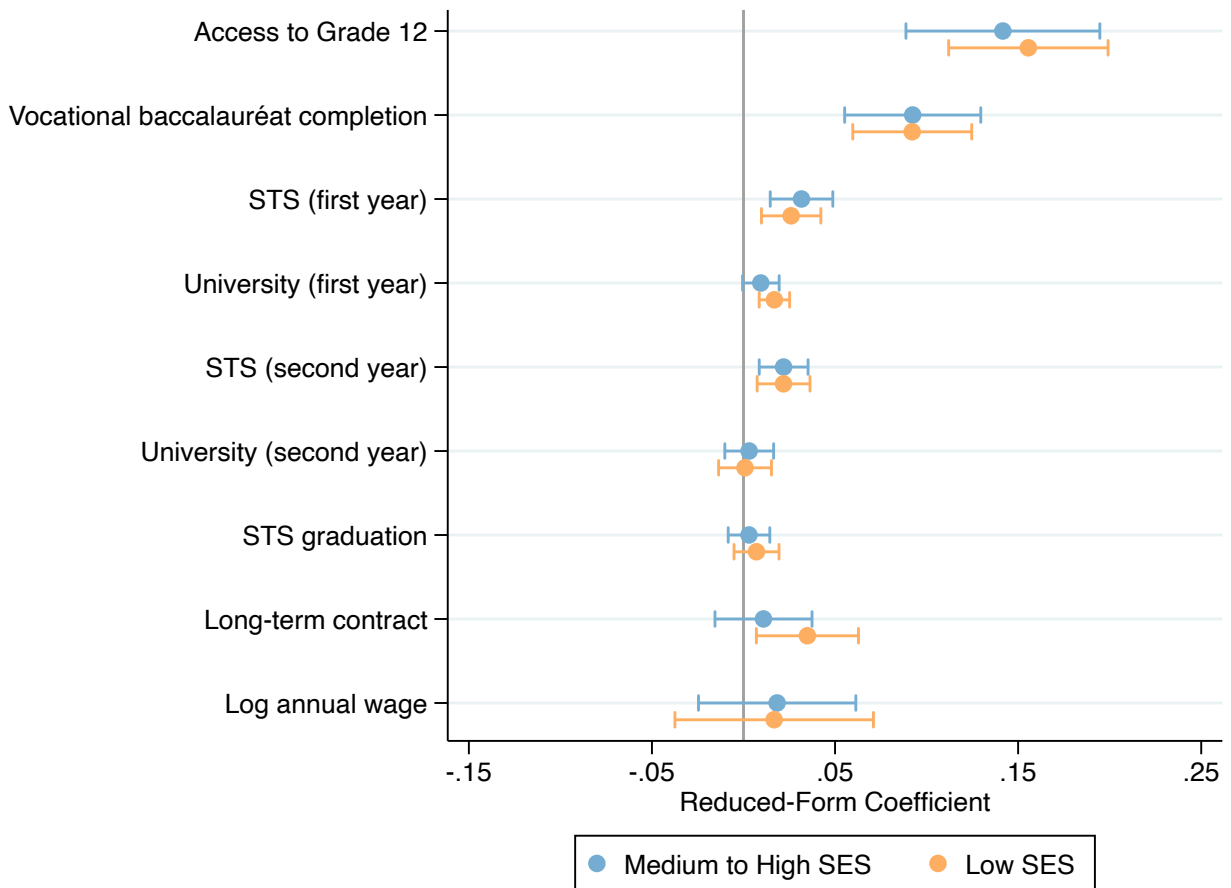
Notes: The figure reports the coefficients on predicted exposure to vocational programs conversions obtained from the linear estimation of Equation (1.1) on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year, and were wage earners in the DADS data 2016–2018. The coefficients are estimated based on administrative data from the French Ministries of Education and Higher Education and Insee.

Figure 1.11 – Reduced-Form Coefficients on Exposure to Vocational Program Conversions: Female vs. Male Students



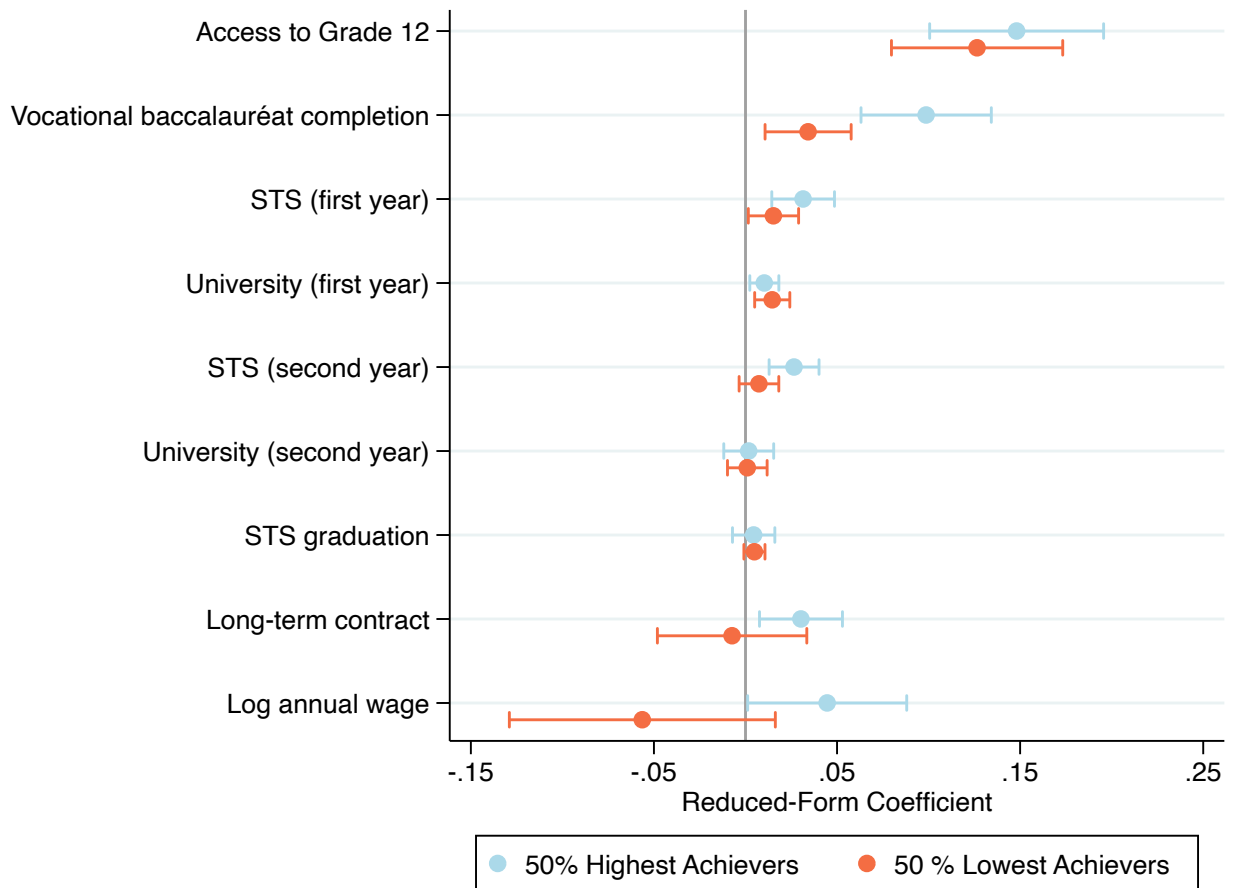
Notes: The figure reports the coefficients on predicted exposure to vocational programs conversions obtained from the linear estimation of Equation (1.1) on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year, and were wage earners in the DADS data 2016–2018. The coefficients are estimated based on administrative data from the French Ministries of Education and Higher Education and Insee.

Figure 1.12 – Reduced-Form Coefficients on Exposure to Vocational Program Conversions: Low vs. Medium-Low to High SES



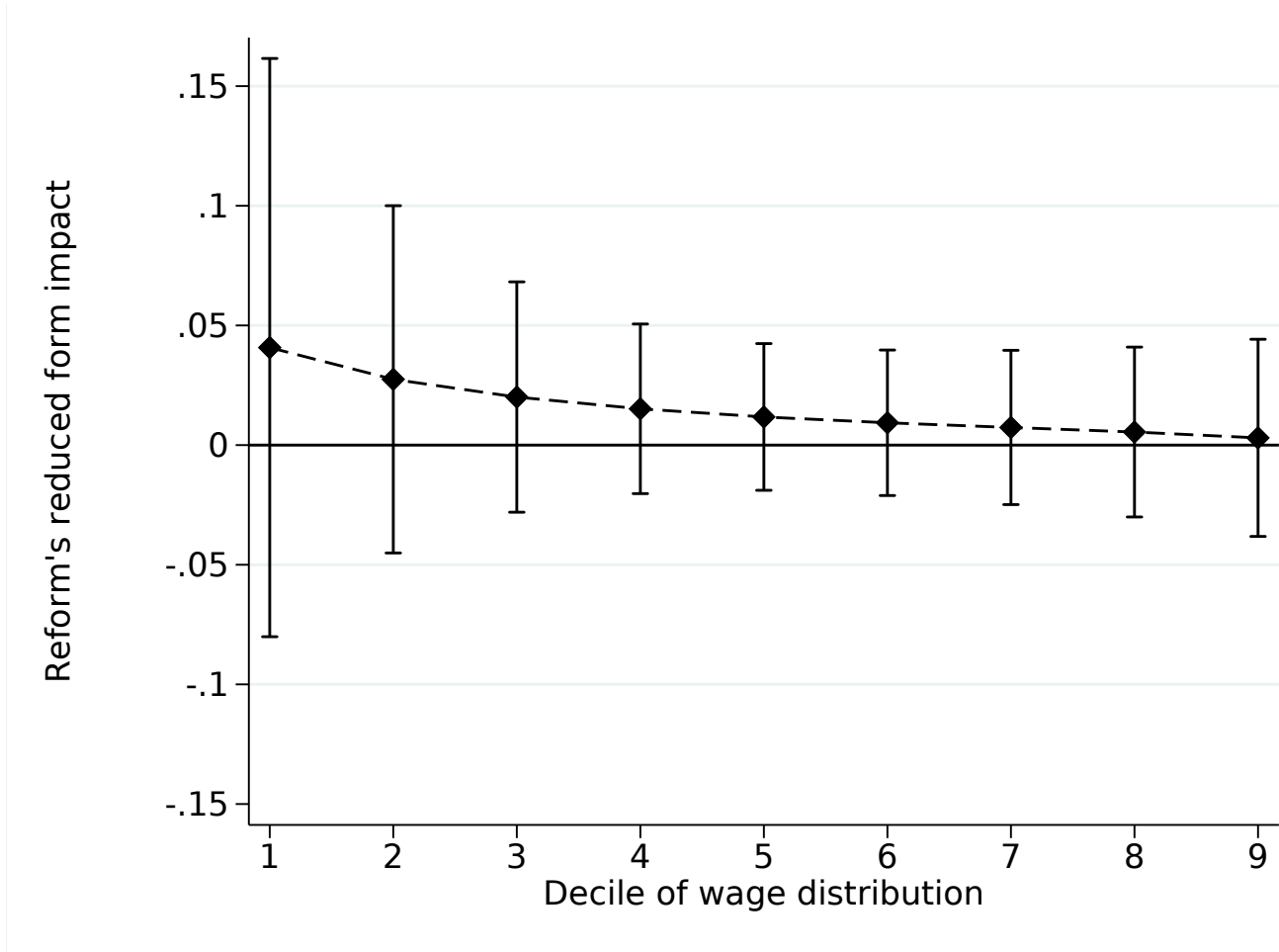
Notes: The figure reports the coefficients on predicted exposure to vocational programs conversions obtained from the linear estimation of Equation (1.1) on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year, and were wage earners in the DADS data 2016–2018. The coefficients are estimated based on administrative data from the French Ministries of Education and Higher Education and Insee.

Figure 1.13 – Reduced-Form Coefficients on Exposure to Vocational Program Conversions: Low vs. High Prior Achievement Level



Notes: The figure reports the coefficients on predicted exposure to vocational programs conversions obtained from the linear estimation of Equation (1.1) on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year, and were wage earners in the DADS data 2016–2018. The coefficients are estimated based on administrative data from the French Ministries of Education and Higher Education and Insee.

Figure 1.14 – Coefficients from the Quantile Regression of Log Annual Wages on Exposure to the Converted Programs



Notes: The figure reports the coefficients on predicted exposure to the reform obtained from the decile regression of Model 1.1 with dependent variable log annual wages nine years after Grade 9 and estimated on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year, and were wage earners in the DADS data 2016–2018. The coefficients are estimated based on administrative data from the French Ministries of Education and Higher Education and the Insee.

Table 1.1 – Odds Ratios of the Probability of BEP Program Conversion in Year 2009, 2010, or 2011 (base 2008)

	Year of conversion (base 2008)		
	2009 (1)	2010 (2)	2011 (3)
Characteristics of BEP students in 2007			
% DNB candidates	0.193*** (0.038)	0.161** (0.070)	0.175*** (0.044)
Average DNB rank (DNB candidates only)	0.982*** (0.003)	1.005*** (0.010)	1.044*** (0.005)
% males	1.724*** (0.135)	0.960*** (0.218)	0.059*** (0.008)
% low SES	3.787*** (0.664)	0.966** (0.492)	1.137*** (0.290)
% accessing 11 th grade	0.081*** (0.013)	0.003** (0.001)	0.032*** (0.007)
% completing vocational <i>baccalauréat</i>	0.077*** (0.013)	0.002* (0.001)	0.014*** (0.004)
Vocational specialization			
Industrial transformation	1.632*** (0.236)	1.104** (0.481)	4.400*** (0.712)
Civil engineering	3.567*** (0.420)	0.934** (0.376)	0.048 (0.034)
Mechanics, electricity	1.093*** (0.068)	0.307*** (0.079)	0.028*** (0.009)
Management	0.463*** (0.029)	0.746*** (0.138)	0.203*** (0.026)
Communication	0.600*** (0.047)	0.595*** (0.154)	0.146*** (0.031)
Care	1.204*** (0.192)	9.608*** (2.248)	61.118*** (9.430)

Notes: For each baseline BEP characteristic listed in the first column, each row reports the odds ratios of BEP conversion in the year listed in column compared to 2008. The estimation is performed on the universe of 2006 to 2011 BEP programs based on administrative data from the French Ministry of Education. Standard errors are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

Table 1.2 – Pre-Reform Shares Assigned by Middle Schools to High Schools

	Mean (1)	S.D. (2)
Middle school level statistics		
Number of high schools assigned non-zero weight	10.517	5.428
Share of high schools assigned non-zero weight	0.005	0.003
Max weight assigned to high school	0.400	0.190
Min weight assigned to high school	0.061	0.124
<hr/>		
N middle schools	8193	
N high schools	1977	

Notes: The table reports summary statistics on the shares assigned by middle schools to high schools that are used to compute the Bartik instrument. The figures are computed on the universe of 2006 to 2007 Grade 9 students enrolling in a vocational program after Grade 9, based on administrative data from the French Ministry of Education.

Table 1.3 – Summary Statistics: Vocational vs. Non-Vocational Students

	All students		Pre-reform vocational students		
	Non vocational (1)	Vocational (2)	CAP (3)	2-year BEP (4)	2+2-year BEP (5)
Prior academic achievement					
DNB candidate	0.880 (0.325)	0.825 (0.380)	0.453 (0.498)	0.878 (0.327)	0.944 (0.231)
DNB rank (DNB candidates only)	56.900 (27.629)	28.553 (22.074)	27.706 (23.636)	25.107 (22.269)	32.236 (21.179)
Socio-demographic characteristics					
Female	0.513 (0.500)	0.438 (0.496)	0.468 (0.499)	0.421 (0.494)	0.452 (0.498)
SES					
High	0.260 (0.438)	0.067 (0.250)	0.046 (0.210)	0.059 (0.236)	0.080 (0.272)
Medium-High	0.153 (0.360)	0.109 (0.312)	0.085 (0.279)	0.104 (0.306)	0.127 (0.333)
Medium-Low	0.269 (0.443)	0.267 (0.442)	0.220 (0.414)	0.255 (0.436)	0.287 (0.452)
Low	0.319 (0.466)	0.557 (0.497)	0.649 (0.477)	0.581 (0.493)	0.506 (0.500)
<hr/>					
N	2,152,456	688,742	60,305	147,836	140,523
Cohorts	2006–2009	2006–2009	2006–2007	2006–2007	2006–2007

Notes: The table reports the average values of the variables listed in the first column together with their standard deviations (in parentheses) for non-vocational vs. vocational track students (Columns 1 and 2) and separately for the vocational students of each type of pre-reform vocational program (Columns 3 to 5). Non-vocational students are general and technological track students, Grade 9 repeaters, dropouts and apprentices. The figures are computed on the universe of students who were enrolled in Grade 9 between 2006 and 2009 and who attended a vocational track program the following year, based on administrative data from the French Ministry of Education.

Table 1.4 – Descriptive Evidence on Reform’s Impact: Vocational vs. non-Vocational Students

	Cohort 2007		Cohort 2009	
	Non-vocational students (1)	Vocational students (2)	Non-vocational students (3)	Vocational students (4)
Outcomes within 4 years				
Access to 12 th grade	0.699	0.361	0.742	0.540
High school graduation	0.618	0.274	0.654	0.374
Outcomes within 5 years				
Access to higher education	0.506	0.109	0.532	0.148
Access to first year STS	0.111	0.076	0.116	0.104
Access to first year university	0.286	0.028	0.306	0.045
Outcomes within 7 years				
STS graduation	0.072	0.038	0.075	0.045
University graduation	0.162	0.010	0.168	0.012
Employment status				
Employed	0.691	0.703	0.724	0.732
Enrolled in education	0.120	0.008	0.108	0.005
Labour market outcomes ^a				
Long-term contract	0.461	0.509	0.474	0.540
Annual wage (2016 euros)	16,548	16,130	16,817	16,483
Hourly wage (2016 euros)	12.867	12.477	13.190	12.750
N	550,326	177,745	532,492	170,735

Notes: ^a Computed for individuals who are not enrolled in education.

The table reports the average values of the outcomes listed in the first column for the last pre-reform cohorts of non-vocational vs. vocational track students (Columns 1 and 2) and the 2009 cohorts of non-vocational vs. vocational track students (Columns 3 and 4). Non-vocational students are general and technological track students, Grade 9 repeaters, dropouts and apprentices. The figures are computed on the universe of students who were enrolled in Grade 9 between 2006 and 2009 and who attended a vocational track program the following year, based on administrative data from the French Ministry of Education and the Insee.

Table 1.5 – Reform’s Reduced-Form Impact: Educational Outcomes

	Baseline mean	Reduced-form estimates			
	(1)	(2)	(3)	(4)	(5)
Number of years of secondary schooling	2.958	−0.335*** (0.038)	−0.345*** (0.034)	−0.313*** (0.043)	−0.316*** (0.038)
Outcomes within 2 years					
High school drop out	0.160	0.023*** (0.007)	0.026*** (0.008)	0.028*** (0.008)	0.029*** (0.008)
Grade retention	0.111	0.036*** (0.012)	0.037*** (0.013)	0.051*** (0.009)	0.052*** (0.009)
Outcomes within 4 years					
Access to 12 th grade	0.355	0.148*** (0.017)	0.143*** (0.018)	0.155*** (0.019)	0.153*** (0.019)
High school graduation	0.265	0.100*** (0.014)	0.094*** (0.015)	0.095*** (0.014)	0.093*** (0.014)
Outcomes within 5 years					
Access to first year of higher education	0.109	0.031*** (0.006)	0.029*** (0.006)	0.028*** (0.007)	0.027*** (0.007)
Access to first year of STS	0.078	0.022*** (0.005)	0.021*** (0.005)	0.019*** (0.006)	0.018*** (0.006)
Access to first year of university	0.026	0.017*** (0.003)	0.016*** (0.003)	0.014*** (0.003)	0.014*** (0.003)
Outcomes within 6 years					
Access to second year of higher education	0.060	0.013*** (0.004)	0.012*** (0.004)	0.011** (0.005)	0.010** (0.005)
Access to second year of STS	0.054	0.014*** (0.004)	0.014*** (0.004)	0.012*** (0.004)	0.011*** (0.004)
Access to second year of university	0.002	0.002** (0.001)	0.002** (0.001)	0.002*** (0.001)	0.002** (0.001)
Degree completion in higher education within 7 years					
STS graduation	0.039	0.009** (0.004)	0.008** (0.004)	0.006 (0.004)	0.006 (0.004)
University graduation	0.010	0.002 (0.002)	0.001 (0.002)	0.003 (0.002)	0.003 (0.002)
ECTS credits accumulated	1.340	−0.073 (0.264)	−0.109 (0.257)	−0.167 (0.289)	−0.179 (0.279)
N	352,287	688,715	688,715	688,715	688,715
Cohorts	2006–2007	2006–2009	2006–2009	2006–2009	2006–2009
Controls					
Middle school fixed effects		×	×	×	×
Cohort fixed effects		×	×	×	×
Individual characteristics			×		×
Middle school trends				×	×

Notes: Each row reports the coefficients on predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column. The regression includes middle school and year fixed effects and is estimated on the universe of students who were enrolled in Grade 9 between 2006 and 2009 and who attended a vocational track program the following year, based on administrative data from the French Ministries of Education and Higher Education. Standard errors are clustered at the middle school \times year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

Table 1.6 – Reform’s Reduced-Form Impact: Educational and Labor Market Outcomes - Wage Earners Only

	Baseline mean	Reduced-form estimates	
	(1)	(2)	(3)
Matriculation in vocational track			
CAP	0.158	0.031** (0.014)	0.033** (0.014)
Seconde professionnelle	0.011	0.857*** (0.023)	0.858*** (0.023)
Educational outcomes			
Number of years of secondary schooling	3.046	-0.336*** (0.037)	-0.344*** (0.036)
Access to 12 th grade	0.401	0.148*** (0.022)	0.145*** (0.022)
High school graduation	0.313	0.090*** (0.015)	0.085*** (0.015)
Dropout	0.142	0.019** (0.008)	0.022*** (0.008)
Access to first year STS	0.091	0.029*** (0.007)	0.028*** (0.007)
Access to first year university	0.028	0.013*** (0.003)	0.012*** (0.003)
Access to second year STS	0.065	0.022*** (0.005)	0.021*** (0.005)
Access to second year university	0.002	0.002 (0.005)	0.002 (0.005)
STS graduation	0.047	0.006 (0.004)	0.005 (0.004)
Labour market outcomes			
Long-term contract	0.507	0.022** (0.009)	0.022** (0.009)
Log annual wage	9.351	0.018 (0.019)	0.024 (0.018)
Log hourly wage	2.495	0.005 (0.005)	0.007 (0.004)
N	124,361	372,658	372,658
Cohorts	2007	2007–2009	2007–2009
Controls			
Middle school fixed effects		×	×
Cohort fixed effects		×	×
Individual characteristics			×

Notes: Each row reports the coefficients on predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column. Labor market outcomes are measured nine years after Grade 9. The regression includes middle school and year fixed effects and is estimated the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year and were wage earners in the DADS data 2016–2018, based on administrative data from the French Ministries of Education and Higher Education and the Insee. Standard errors are clustered at the middle school \times year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

Appendix to

Does Time to Graduation Affect Wages?
Evidence from the French Vocational Track Reform.

1.A Data and Sample Restrictions

1.A.1 Data

I analyze students' trajectories in secondary and higher education by combining administrative student registers in secondary and higher education with individual results at national examinations. In addition, for the first time with French administrative data, I was able to match administrative educational data with administrative labor market databases that cover all public and private wage earners living or working in France. The two datasets are merged using a nominative statistical matching procedure that uniquely matches individuals based on their names, first names, gender, and date and place of birth.

Secondary education students. Students' enrollment status in secondary education for the years 2006 to 2017 is derived from the SYSCA administrative datasets of the *Direction de l'Évaluation de la Prospective et de la Performance* (DEPP, French Ministry of Education). The SYSCA databases cover the universe of students enrolled in public or contractual private middle schools and high schools outside apprenticeship programs.^{A.1} Along with basic socioeconomic characteristics (date and place of birth, gender, socioeconomic status of the parents), the SYSCA data provide detailed information on the students' school, Grade level, and program of matriculation.

Higher education students. Students' enrollment status in higher education for the years 2006 to 2017 is derived from the SYSCA databases and the SISE administrative records of the *Sous-direction des Systèmes d'Information et Études statistiques de l'Enseignement Supérieur* (SIES) combined. While the SYSCA databases cover all higher education students enrolled in a program located in high school (CPGE and STS), the SISE databases cover all students enrolled in a university program or in an engineering, business, private or elite school. The combined SYSCA and SISE databases cover almost 90 percent of the students enrolled in higher education each year (MENESR, 2017).^{A.2}

^{A.1}In 2017, 26,879 students were enrolled in a non-contractual private middle school or high school, out of a total of 5,629,800 enrolled in secondary education (MENESR, 2017). The same year, 261,192 out of 2,287,500 high school students were undertaking an apprenticeship program (MENESR, 2017).

^{A.2}Importantly, students enrolled in paramedical higher education programs (5 percent of the students enrolled in higher education each year) are not covered by the data.

Labor market outcomes. Information on students' labor market outcomes between 2016 and 2018 is retrieved from the 2017 and 2018 Insee DADS (*Déclaration annuelle de données sociales*) databases. These administrative databases cover the entire population of public and private wage earners working or living in France in the considered and the preceding year. For every wage earner, the data contain information on the different jobs worked, the number of hours worked and the gross and net wages earned. For each job, a qualitative description is also available, with information on the type of contract and the socioeconomic category of the job.

Database construction. All the information contained in the educational databases can be gathered into one single educational database thanks to an encrypted version of the students' national identifier. This comprehensive educational database consists of the 6,576,000 students who were enrolled in Grade 9 between 2006 and 2014, of which 1,608,072 joined the vocational track just after Grade 9.

The labor market database is matched to the educational database using a statistical matching procedure based on the first and last names, the gender, and the date and place of birth of the students. Table 1.A1 reports the rates of wage earners uniquely found in the educational database and the rate of students uniquely matched in the educational database, by year of birth.^{A.3} Among the wage earners born between 1990 to 1999, between 85.6 percent and 92.4 percent can be uniquely identified in the educational database. Symmetrically, the rate of students that can be uniquely identified in the 2016 to 2018 wage data is higher than 75 percent for the students born in the early 90s, and decreases sharply for the 1998 and 1999 birth cohorts, reflecting the fact that a non-negligible share of individuals from these cohorts are still enrolled in education between 2016 and 2018.

^{A.3}The rates are computed among French students and French wage earners only.

Table 1.A1 – Statistical Match Between Educational and Labor Market Data

Birth Cohort	Share students uniquely matched in 2016 to 2018 wage data (1)	Share 2016/2017 wage earners uniquely matched in educational data (2)	Share 2018 wage earners uniquely matched in educational data (3)
1990	0.745	0.856	0.859
1991	0.759	0.902	0.905
1992	0.773	0.912	0.916
1993	0.785	0.915	0.919
1994	0.795	0.917	0.921
1995	0.801	0.917	0.922
1996	0.793	0.917	0.921
1997	0.766	0.914	0.918
1998	0.715	0.922	0.924
1999	0.602	0.910	0.920

Notes: The figures are computed using the universe of the 2006 to 2014 French Grade 9 students and 2016 to 2018 public and private wage earners born in France, using administrative data from the French Ministry of Education and the Insee.

1.A.2 Indicators on students’ socioeconomic and academic profile

Students’ socioeconomic status (SES) is broken down into different categories, using the French Ministry of Education’s official classification. This classification uses the occupation of the child’s legal guardian to define four groups of SES: high (company managers, executives, liberal professions, engineers, intellectual occupations, arts professions), medium-high (technicians and associate professionals), medium-low (farmers, craft and trades workers, service and sales workers), and low (manual workers and persons without employment).

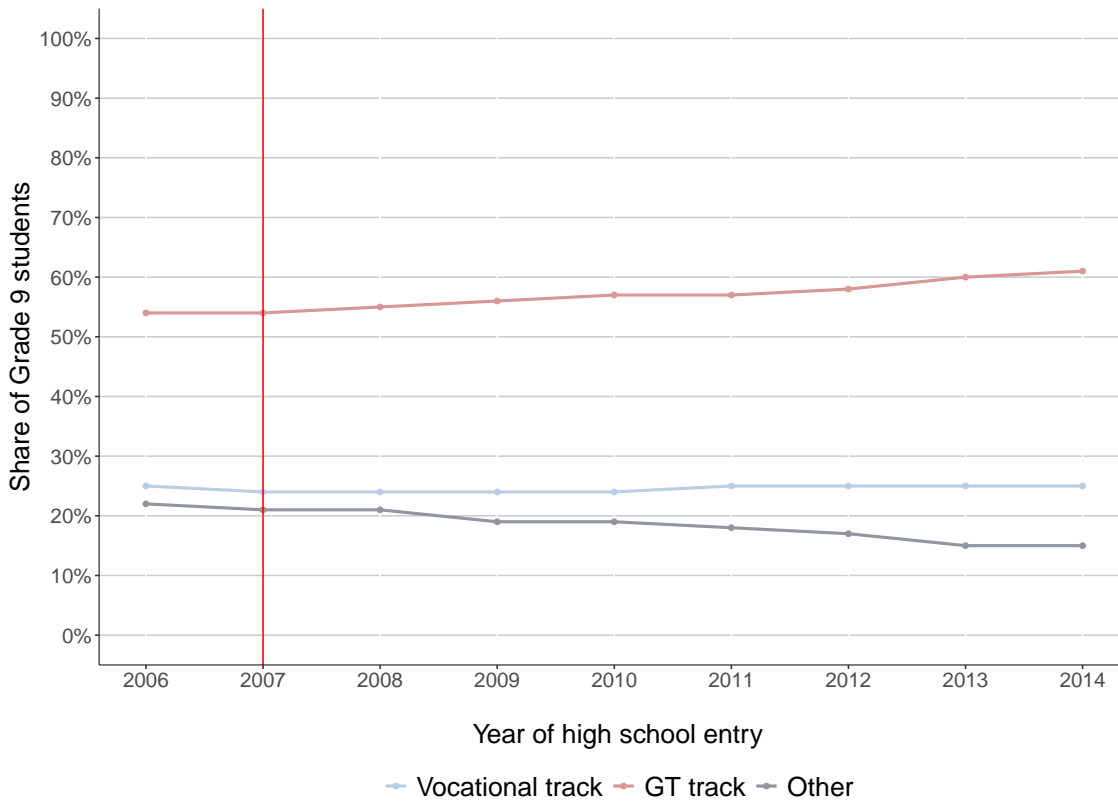
Students’ achievement level prior to entering the vocational track is measured based on the average of their percentile ranks at the written French and math DNB examinations. The students who did not take the DNB examinations are tagged as “non DNB candidates”, a strong signal of low achievement level at the end of middle school.

1.A.3 Sample Restrictions

Sample restriction to vocational track students. To provide evidence in support of the absence of reform-induced composition effects among vocational track students, Figure 1.A1 plots the shares of Grade 9 students attending either the general and technological track, or the vocational track, or repeating Grade 9/dropping out/joining an apprenticeship program in the year following Grade 9, over the 2006–2014 period. The figure shows that the share of students attending the vocational track remained stable over the period. I further test for the absence of composition

effect by estimating Equation (1.1) on vocational track students for a set baseline characteristics including gender, SES, and prior achievement level. Table 1.A2 displays the corresponding estimates. Reassuringly, the reform did not alter most of vocational track students' baseline characteristics. In particular, the average prior academic achievement level and the social background of the students are not significantly different for the students who were fully exposed to the reform compared to those who were unexposed. None the coefficients on predicted exposure to the converted programs are statistically different from 0, except for the coefficient on male students.

Figure 1.A1 – Student Tracking after Grade 9



Notes: The figure plots the share of Grade 9 students matriculating in the vocational track, the general and technological track or none of the two, by theoretical year of high school entry. The Other category includes Grade 9 repeaters, apprentices and dropouts. The figures are computed based on 2006 to 2014 administrative data from the French Ministry of Education.

Table 1.A2 – Reform’s Reduced-Form Impact on Vocational Students’ Characteristics

	All students			Wage earners		
	Baseline Mean	Reduced-form estimates		Baseline Mean	Reduced-form estimates	
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.558	-0.028** (0.012)	-0.026** (0.011)	0.584	-0.032** (0.013)	-0.031** (0.012)
DNB candidate	0.838	0.011 (0.013)	0.010 (0.013)	0.851	0.006 (0.009)	0.005 (0.009)
DNB rank - <i>DNB candidates only</i>	29.084	0.786 (0.735)	0.717 (0.734)	29.123	0.680 (0.468)	0.582 (0.475)
SES						
Low	0.561	-0.001 (0.009)	-0.001 (0.009)	0.542	-0.000 (0.009)	0.000 (0.009)
Average	0.261	-0.000 (0.008)	-0.001 (0.008)	0.272	-0.001 (0.008)	-0.002 (0.008)
High	0.112	0.002 (0.004)	0.001 (0.004)	0.116	0.001 (0.005)	0.001 (0.005)
Very high	0.066	-0.000 (0.003)	0.000 (0.003)	0.069	0.001 (0.004)	0.001 (0.004)
Appear in wage database	0.791	-0.010 (0.008)	-0.010 (0.007)	1.000	-	-
N	178,008	688,715	688,715	124,361	372,658	372,658
Cohorts	2007	2007–2009	2007–2009	2007	2007–2009	2007–2009
Controls						
Middle school fixed effects		x	x		x	x
Cohort fixed effects		x	x		x	x

Notes: Each row reports the coefficients on predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column. Labor market outcomes are measured nine years after Grade 9. The regression includes middle school and year fixed effects and is estimated the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year, based on administrative data from the French Ministries of Education and Higher Education and the Insee. Standard errors are clustered at the middle school \times year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

Wage earners sample restrictions. Table 1.A2 also gives evidence on the absence of compositional effects of the reform on the wage earners subsample of vocational track students. When estimating Equation (1.1) for the probability of appearance in the wage database on the exhaustive sample of vocational students from the cohorts 2007–2009, I find that full exposure to the reform is associated to a 0.01 percentage point decrease in the probability of appearing in the wage database, a coefficient that is statistically insignificant at all conventional levels (see Columns 2 and 3). I can therefore rule out the existence of selection into the wage data induced by the reform.

The table brings additional evidence supporting the absence of compositional effects of the reform on the wage earners subsample of vocational track students. Columns 5 and 6 of the table report the coefficients of the estimation of Equation (1.1) on the wage earners subsample, for a set of baseline students’ characteristics that includes gender, prior achievement level and SES. None of the coefficients on baseline covariates is significantly different from 0 at any conventional level, except the coefficient on being a male student. Full exposure to the reform is predicted to decrease

the share of male students among the wage earners of the vocational track by 3.2 percentage points, with a confidence level of 5 percent. This finding is not induced by the wage earners restriction, as it is prevalent on the exhaustive vocational track students sample (see Columns 2 and 3). I address the issue in Section 1.4.5.

1.B Reform’s Impact on Grade Retention and High School Dropout

Methodological issues. Assessing the impact of the reform on high school dropout and grade retention is complicated by the time inconsistency in the definitions of both grade retention and high school dropout induced by the reform.

During the years of implementation of the reform, an untreated year t BEP student repeating its first or second year of BEP could either repeat in the same program, if the program was not converted between years t and $t + 1$, or in a converted program if the program was converted between t and $t + 1$. While grade repetition is clearly defined in the first case, it is not in the second case, since, by design, the grade that is retaken is in fact different from the first one undertaken by the student. The vocational Grades 11 and 12 were impacted by the reorganization of a four-year schedule into a three-year schedule, yet their status was not affected by the reform from an institutional point of view. I therefore define both the first year and the second year of BEP as being equivalent to a *seconde professionnelle*, i.e., a Grade 10 program.

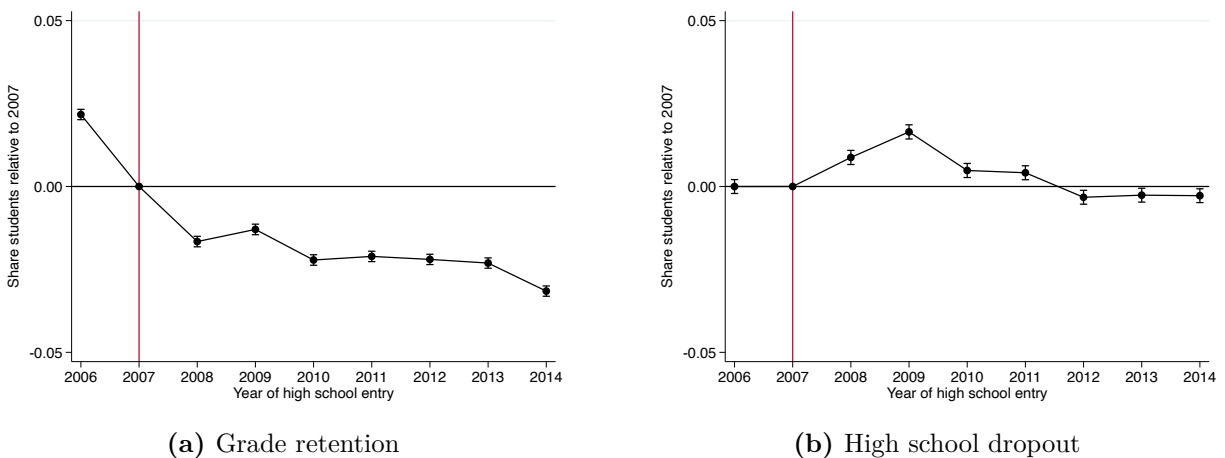
Similarly, the definition of high school dropout poses a number of issues. Defined *stricto sensu*, high school dropout is the non-completion of a high school program that a student started. However, as the reform converted the 2+2-year BEP path into a 3-year path, this definition implies that after the reform, all non-CAP students who drop out between the second and the third year would be considered as dropouts, while, before the reform, only those who would attempt the Grade 11+12 path but fail to complete it would be considered as dropout. To overcome this issue, I consider as high school dropouts only the students who do not complete the two first years of a vocational program. Note that this solution is not perfect either, as it implies that after the reform, failure to access Grade 11 is considered as high school dropout for non-CAP students, while this is not the case for pre-reform students.

Interpretation. Grade retention in the vocational track has been on a decreasing trend throughout the years of implementation of the reform, without any clear link with the reform (see Figure 1.B1). Coefficients from the estimation of Equation (1.1) however suggest a positive effect of full exposure to the reform, which is an intuitive result. In areas that were exposed to early BEP conversions, grade retention is likely to have been perceived as a zero-cost policy for all students wishing to reach the *baccalauréat*, as repeating a grade implied obtaining the *baccalauréat* in four

years rather than three, which was most likely still perceived as the dominant norm. Conversely, in areas that were little exposed to early BEP conversions, grade retention is likely to have been perceived as costly, since, in a non-negligible share of cases, a non-treated grade repeater would repeat in a converted program, for which she likely did not have sufficient academic preparation.

Unlike for grade repetition, the reform had clear effects on the high school dropout rate for vocational track students in the years of implementation of the reform, with a transitory increase that culminated at two percentage points between 2007 and 2009. The dropout rate however gradually decreased until its pre-reform level between the years 2009 and 2011. These findings are consistent with the discussion of Section 1.2 relative to the vocational track teachers' learning process that took place in the years of implementation of the reform.

Figure 1.B1 – Reform's Reduced-Form Impact on Grade Retention and High School Dropout



Notes: The figure displays the share of vocational track students repeating a grade or dropping out from high school, by year of high school entry. The figures are computed on the universe of students who were enrolled in Grade 9 between 2006 and 2014 and who attended a vocational track program the following year, based on administrative data from the French Ministries of Education and Higher Education. The vertical red line denotes the last pre-reform year.

1.C Heterogeneity analysis

Table 1.C1 – Reform’s Reduced-Form Impact on Educational and Labor Market Outcomes, by Gender - Wage Earners Only

	Female		Male	
	Baseline (1)	Estimate (2)	Baseline (3)	Estimate (4)
Matriculation in vocational track				
CAP	0.166	0.029 (0.019)	0.152	0.031* (0.017)
Seconde professionnelle	0.014	0.817*** (0.029)	0.010	0.871*** (0.028)
Educational outcomes				
Number of years of secondary schooling	3.087	−0.359*** (0.043)	3.016	−0.323*** (0.046)
Access to 12 th grade	0.419	0.139*** (0.029)	0.388	0.149*** (0.024)
High school graduation	0.332	0.086*** (0.022)	0.299	0.087*** (0.017)
Dropout	0.119	0.010 (0.011)	0.158	0.024** (0.010)
Access to first year STS	0.088	0.030*** (0.011)	0.092	0.029*** (0.008)
Access to first year university	0.039	0.023*** (0.006)	0.021	0.007* (0.004)
Access to second year STS	0.062	0.028*** (0.009)	0.067	0.019*** (0.006)
Access to second year university	0.003	0.004 (0.009)	0.002	0.001 (0.006)
STS graduation	0.044	0.006 (0.007)	0.049	0.006 (0.006)
Labour market outcomes				
Long-term contract	0.493	0.005 (0.013)	0.517	0.033*** (0.011)
Log annual wage	9.185	−0.008 (0.032)	9.470	0.054** (0.021)
Log hourly wage	2.451	0.002 (0.006)	2.525	0.012** (0.005)
N	51,080	153,677	70,703	215,731
Cohorts	2007	2007–2009	2007	2007–2009
Controls				
Middle school fixed effects		×		×
Cohort fixed effects		×		×

Notes: Each row reports the coefficients on predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column, estimated separately for female and male students. Labor market outcomes are measured nine years after Grade 9. The regression includes middle school and year fixed effects and is estimated on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year and were wage earners in the DADS data 2016–2018, based on administrative data from the French Ministries of Education and Higher Education and the Insee. Standard errors are clustered at the middle school \times year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

Table 1.C2 – Reform’s Reduced-Form Impact on Educational and Labor Market Outcomes, by SES - Wage Earners Only

	Medium-Low to High SES		Low SES	
	Baseline (1)	Estimate (2)	Baseline (3)	Estimate (4)
Matriculation in vocational track				
CAP	0.133	0.028* (0.017)	0.180	0.033** (0.017)
Seconde professionnelle	0.013	0.844*** (0.029)	0.010	0.872*** (0.024)
Educational outcomes				
Number of years of secondary schooling	3.130	−0.367*** (0.040)	2.975	−0.316*** (0.041)
Access to 12 th grade	0.449	0.142*** (0.027)	0.360	0.156*** (0.022)
High school graduation	0.356	0.092*** (0.019)	0.277	0.092*** (0.017)
Dropout	0.122	0.018* (0.010)	0.158	0.023** (0.009)
Access to first year STS	0.106	0.032*** (0.009)	0.077	0.026*** (0.008)
Access to first year university	0.027	0.009* (0.005)	0.029	0.017*** (0.004)
Access to second year STS	0.078	0.022*** (0.007)	0.053	0.022*** (0.007)
Access to second year university	0.003	0.003 (0.007)	0.002	0.001 (0.007)
STS graduation	0.057	0.003 (0.006)	0.038	0.007 (0.006)
Labour market outcomes				
Long-term contract	0.533	0.011 (0.014)	0.485	0.035** (0.014)
Log annual wage	9.434	0.018 (0.022)	9.281	0.017 (0.028)
Log hourly wage	2.504	0.002 (0.006)	2.487	0.011* (0.006)
N	55,734	170,662	66,049	198,710
Cohorts	2007	2007–2009	2007	2007–2009
Controls				
Middle school fixed effects		×		×
Cohort fixed effects		×		×

Notes: Each row reports the coefficients on predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column, estimated separately for low and medium-low to high SES students. Labor market outcomes are measured nine years after Grade 9. The regression includes middle school and year fixed effects and is estimated on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year and were wage earners in the DADS data 2016–2018, based on administrative data from the French Ministries of Education and Higher Education and the Insee. Standard errors are clustered at the middle school × year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

Table 1.C3 – Reform’s Reduced-Form Impact on Educational and Labor Market Outcomes, by Prior Achievement Level - Wage Earners Only

	50 % Highest Achievers		50 % Lowest Achievers	
	Baseline (1)	Estimate (2)	Baseline (3)	Estimate (4)
Matriculation in vocational track				
CAP	0.086	0.023** (0.011)	0.296	0.081*** (0.026)
Seconde professionnelle	0.014	0.879*** (0.029)	0.006	0.763*** (0.023)
Educational outcomes				
Number of years of secondary schooling	3.222	−0.391*** (0.035)	2.710	−0.219*** (0.045)
Access to 12 th grade	0.491	0.148*** (0.024)	0.228	0.126*** (0.024)
High school graduation	0.407	0.099*** (0.018)	0.133	0.034*** (0.012)
Dropout	0.097	0.022** (0.009)	0.227	0.030** (0.012)
Access to first year STS	0.118	0.031*** (0.009)	0.037	0.015** (0.007)
Access to first year university	0.035	0.010** (0.004)	0.015	0.015*** (0.005)
Access to second year STS	0.086	0.027*** (0.007)	0.024	0.007 (0.006)
Access to second year university	0.003	0.002 (0.007)	0.001	0.001 (0.006)
STS graduation	0.065	0.004 (0.006)	0.012	0.005* (0.003)
Labour market outcomes				
Long-term contract	0.535	0.030*** (0.012)	0.454	−0.007 (0.021)
Log annual wage	9.454	0.045** (0.022)	9.155	−0.056 (0.037)
Log hourly wage	2.505	0.007 (0.005)	2.474	0.003 (0.008)
N	79,875	239,220	41,908	129,970
Cohorts	2007	2007–2009	2007	2007–2009
Controls				
Middle school fixed effects		×		×
Cohort fixed effects		×		×

Notes: Each row reports the coefficients on predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column, estimated separately for the 50 percent highest achievers and the 50 percent lowest achievers of the vocational track. Labor market outcomes are measured nine years after Grade 9. The regression includes middle school and year fixed effects and is estimated on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year and were wage earners in the DADS data 2016–2018, based on administrative data from the French Ministries of Education and Higher Education and the Insee. Standard errors are clustered at the middle school \times year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

1.D Robustness checks

1.D.1 Wald-TC Estimator

As shown by de Chaisemartin and d’Haultfoeuille (2018) and de Chaisemartin and d’Haultfoeuille (2019), in a multiple groups and periods difference-in-differences, the linear estimator computes a weighted average of the average treatment effects of each middle school and year, with weights that can be negative for some middle schools and years, therefore leading to a downward bias in the coefficient of interest. This can happen, in particular, in the presence of dynamic treatment effects.

The Wald-TC estimator allows to relax the common trend assumption on which the traditional linear estimation of the difference-in-differences relies in order to allow for dynamic treatment effects. The estimator requires the existence of a common trend between $t - 1$ and t for the middle schools that had the same level of exposure to the treatment in $t - 1$. The estimator compares the evolution in the outcomes of middle schools with exposure level $Exp\hat{posure}_{m,t-1} = e$, for middle schools whose exposure level increased between $t - 1$ and t compared to middle schools whose exposure level remained constant ($Exp\hat{posure}_{m,t-1} = Exp\hat{posure}_{m,t} = e$). The Wald-TC estimator therefore requires the existence of a “common support” at each exposure level e : for each e , there needs to exist a sufficiently large control group, i.e. there needs to be enough middle schools whose exposure level did not increase.

When the considered treatment is continuous, several parametrizations have to be done before performing the estimation. With a continuous exposure level that takes a very large number of values, the common support condition cannot be verified at each exposure level. I therefore classify middle schools into quintiles of predicted exposure to the reform. In addition, the continuous nature of the instrument imposes to define a threshold value under which the middle schools can be considered as untreated and therefore part of the control group. This threshold value has to be sufficiently high so that the control group is not too small, and sufficiently small so that the corresponding variation in exposure to the reform is unlikely to affect these middle schools’ outcomes.

Table 1.D1 reports the coefficients obtained under different parameter values for the control group threshold. Between 2008 and 2009, virtually all middle schools’ exposure level increased by a sizable amount. As a consequence, for the estimator to exploit the 2008/2009 variation in the data (which is verified when the control group is not of size 0), the threshold value has to be set at a high level. The cost of this parametrization is that a large number of highly treated middle schools

have to be assigned to the control group. Therefore, the setting of the reform, by constraining the parametrization choices, implies that the Wald-TC approach reintroduces a downward bias in the estimation. Unsurprisingly, the coefficients obtained are therefore similar to the main specification's coefficients.

Table 1.D1 – Reform’s Reduced-Form Impact - Wald-TC Estimator

	Reduced-form estimates		
	(1)	(2)	(3)
Seconde professionnelle	0.856*** (0.057)	0.955*** (0.029)	0.961*** (0.026)
CAP	0.019 (0.050)	-0.006 (0.032)	-0.022 (0.033)
Number of years of secondary schooling	-0.292** (0.118)	-0.337*** (0.075)	-0.352*** (0.088)
Access to Grade 12	0.185*** (0.056)	0.207*** (0.026)	0.172*** (0.032)
Vocational <i>baccalauréat</i> completion	0.086** (0.036)	0.117*** (0.028)	0.089*** (0.023)
Dropout	0.015 (0.058)	-0.072** (0.034)	-0.051* (0.027)
Access to first year of STS	0.051 (0.035)	0.024 (0.016)	0.016 (0.013)
Access to first year of university	0.034*** (0.009)	0.021** (0.009)	0.024*** (0.008)
Access to second year of STS	0.045** (0.022)	0.017 (0.013)	0.010 (0.012)
Access to second year of university	0.005* (0.003)	0.006 (0.006)	0.003 (0.005)
Long-term contract	-0.035 (0.065)	-0.058 (0.046)	-0.028 (0.043)
Log annual wage	0.173 (0.139)	0.074 (0.096)	0.046 (0.095)
Log hourly wage	0.024 (0.020)	0.020 (0.016)	0.020 (0.014)
N obs used for identification	243,146	243,681	243,676
of which: N switchers	208,830	188,445	165,154
Control middle schools: change in exposure lower than:	10%	20%	30%

Notes: Each row reports the coefficients on predicted exposure to the reform from the Wald-TC estimation for the different outcomes listed in the first column. The regression is estimated on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year and were wage earners in the DADS data from 2016–2018 based on administrative data from the French Ministries of Education and Higher Education and from the Insee. Standard errors are clustered at the middle school \times year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05; *: p-value < 0.1.

1.D.2 Extending the Analysis to Non-Vocational Students

Table 1.D2 – Reform’s Reduced-Form Impact - Vocational and Non-Vocational Students, Wage Earners Only

	Baseline mean	Reduced-form estimates	
	(1)	(2)	(3)
Matriculation in vocational track			
CAP	0.046	0.055*** (0.007)	0.055*** (0.007)
Seconde professionnelle	0.003	0.527*** (0.027)	0.528*** (0.028)
Educational outcomes			
Number of years of secondary schooling	3.225	−0.075** (0.036)	−0.076** (0.036)
Access to 12 th grade	0.721	0.153*** (0.015)	0.151*** (0.015)
High school graduation	0.630	0.089*** (0.011)	0.087*** (0.010)
Dropout	0.093	−0.009 (0.009)	−0.008 (0.009)
Access to first year STS	0.135	0.035*** (0.005)	0.035*** (0.005)
Access to first year university	0.238	−0.002 (0.007)	−0.003 (0.007)
Access to second year STS	0.111	0.024*** (0.004)	0.024*** (0.004)
Access to second year university	0.130	−0.014*** (0.004)	−0.015*** (0.004)
STS graduation	0.111	0.010*** (0.004)	0.010*** (0.004)
Labour market outcomes			
Long-term contract	0.439	0.019*** (0.006)	0.020*** (0.006)
Log annual wage	9.281	−0.001 (0.018)	−0.001 (0.018)
Log hourly wage	2.487	0.002 (0.007)	0.002 (0.007)
N	423,975	1,295,424	1,295,424
Cohorts	2007	2007–2009	2007–2009
Controls			
Middle school fixed effects		×	×
Cohort fixed effects		×	×
Individual characteristics			×

Notes: Each row reports the coefficients on predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column. Labor market outcomes are measured nine years after Grade 9. The regression includes middle school and year fixed effects and is estimated on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and were wage earners in the DADS data from 2016–2018, based on administrative data from the French Ministries of Education and Higher Education and the Insee. Standard errors are clustered at the middle school × year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

1.D.3 Exploiting Variation at the *Département* and *Académie* Levels

Table 1.D3 – Reform’s Reduced-Form Impact Using Variation at the *Département* Level - Wage Earners Only

	Baseline mean	Reduced-form estimates	
	(1)	(2)	(3)
Matriculation in vocational track			
CAP	0.158	0.068*** (0.025)	0.072*** (0.024)
Seconde professionnelle	0.011	0.987*** (0.040)	0.988*** (0.040)
Educational outcomes			
Number of years of secondary schooling	3.046	−0.363*** (0.065)	−0.373*** (0.063)
Access to 12 th grade	0.401	0.179*** (0.039)	0.175*** (0.038)
High school graduation	0.313	0.100*** (0.023)	0.094*** (0.022)
Dropout	0.142	0.029** (0.013)	0.033** (0.014)
Access to first year STS	0.091	0.032*** (0.011)	0.031*** (0.011)
Access to first year university	0.028	0.020*** (0.005)	0.019*** (0.005)
Access to second year STS	0.065	0.025*** (0.009)	0.024*** (0.008)
Access to second year university	0.002	0.003 (0.009)	0.003 (0.008)
STS graduation	0.047	0.010 (0.008)	0.009 (0.007)
Labour market outcomes			
Long-term contract	0.507	0.012 (0.016)	0.012 (0.016)
Log annual wage	9.351	−0.024 (0.026)	−0.017 (0.024)
Log hourly wage	2.495	0.001 (0.008)	0.003 (0.007)
N	124,361	372,693	372,693
Cohorts	2007	2007–2009	2007–2009
Controls			
<i>Département</i> fixed effects		×	×
Cohort fixed effects		×	×
Individual characteristics			×

Notes: Each row reports the coefficients on the *départements*’ predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column. Labor market outcomes are measured nine years after Grade 9. The regression includes *département* and year fixed effects and is estimated on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year and were wage earners in the DADS data 2016–2018, based on administrative data from the French Ministries of Education and Higher Education and the Insee. Standard errors are clustered at the *département* × year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

Table 1.D4 – Reform’s Reduced-Form Impact Using Variation at the *Académie* Level- Wage Earners Only

	Baseline mean	Reduced-form estimates	
	(1)	(2)	(3)
Matriculation in vocational track			
CAP	0.158	0.088*** (0.028)	0.089*** (0.027)
Seconde professionnelle	0.011	1.016*** (0.047)	1.019*** (0.046)
Educational outcomes			
Number of years of secondary schooling	3.046	−0.374*** (0.072)	−0.382*** (0.070)
Access to 12 th grade	0.401	0.185*** (0.045)	0.182*** (0.044)
High school graduation	0.313	0.103*** (0.026)	0.099*** (0.026)
Dropout	0.142	0.030* (0.016)	0.034** (0.016)
Access to first year STS	0.091	0.032** (0.012)	0.031** (0.012)
Access to first year university	0.028	0.016*** (0.006)	0.015*** (0.006)
Access to second year STS	0.065	0.026*** (0.009)	0.025*** (0.009)
Access to second year university	0.002	0.002 (0.009)	0.001 (0.009)
STS graduation	0.047	0.010 (0.008)	0.010 (0.008)
Labour market outcomes			
Long-term contract	0.507	0.011 (0.016)	0.012 (0.016)
Log annual wage	9.351	−0.026 (0.028)	−0.014 (0.025)
Log hourly wage	2.495	−0.002 (0.008)	0.001 (0.008)
N	124,361	372,693	372,693
Cohorts	2007	2007–2009	2007–2009
Controls			
<i>Académie</i> fixed effects		×	×
Cohort fixed effects		×	×
Individual characteristics			×

Notes: Each row reports the coefficients on the *académies*’ predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column. Labor market outcomes are measured nine years after Grade 9. The regression includes *académie* and year fixed effects and is estimated on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year and were wage earners in the DADS data 2016–2018, based on administrative data from the French Ministries of Education and Higher Education and the Insee. Standard errors are clustered at the *académie* × year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

1.E Mechanisms investigation

1.E.1 Potential Mechanisms

Table 1.E1 – Reform’s Potential Mechanisms - Wage Earners Only

	Baseline mean	Reduced-form estimates	
	(1)	(2)	(3)
Matriculation in vocational track			
Seconde professionnelle	0.011	0.857*** (0.023)	0.858*** (0.023)
CAP	0.158	0.031** (0.014)	0.033** (0.014)
Grade 11	0.479	0.116*** (0.021)	0.113*** (0.020)
Class size	13.681	-1.938*** (0.424)	-2.044*** (0.439)
Vocational specialization			
Industrial engineering	0.283	-0.025* (0.013)	-0.011 (0.010)
Agriculture and natural landscapes	0.002	0.001 (0.001)	0.001 (0.001)
Industrial transformation	0.058	-0.004 (0.005)	-0.003 (0.004)
Civil engineering	0.097	-0.013** (0.005)	-0.008* (0.005)
Soft materials	0.026	0.006* (0.003)	0.004 (0.003)
Management	0.243	0.015 (0.016)	0.011 (0.015)
Communication	0.096	0.010 (0.014)	0.005 (0.013)
Care	0.176	0.016 (0.010)	0.006 (0.012)
N	124,361	372,658	372,658
Cohorts	2007	2007–2009	2007–2009
Controls			
Middle school fixed effects		×	×
Cohort fixed effects		×	×
Individual characteristics			×

Notes: Each row reports the coefficients on predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column. The regression includes middle school and year fixed effects and is estimated on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year, based on administrative data from the French Ministries of Education and Higher Education. Standard errors are clustered at the middle school \times year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

1.E.2 Excluding the CAP Matriculation and Class Size Channels

Table 1.E2 – Reform’s Reduced-Form Impact on Vocational Students’ Outcomes - Wage Earners Only, *Départements* without CAP effect

	Baseline mean	Reduced-form estimates	
	(1)	(2)	(3)
Matriculation in vocational track			
CAP	0.160	0.012 (0.013)	0.018 (0.012)
Seconde professionnelle	0.013	0.853*** (0.024)	0.853*** (0.024)
Class size	13.777	-1.849*** (0.443)	-1.993*** (0.467)
Educational outcomes			
Access to 12 th grade	0.402	0.157*** (0.024)	0.152*** (0.024)
High school graduation	0.313	0.097*** (0.016)	0.091*** (0.015)
Dropout	0.141	0.011 (0.009)	0.015* (0.009)
Access to first year STS	0.093	0.032*** (0.007)	0.031*** (0.007)
Access to first year university	0.028	0.013*** (0.003)	0.012*** (0.004)
Access to second year STS	0.066	0.026*** (0.006)	0.026*** (0.005)
Access to second year university	0.002	0.002 (0.006)	0.002 (0.005)
STS graduation	0.048	0.009* (0.005)	0.008* (0.005)
Labour market outcomes			
Long-term contract	0.505	0.030*** (0.009)	0.029*** (0.009)
Log annual wage	9.348	0.016 (0.019)	0.021 (0.019)
Log hourly wage	2.494	0.007 (0.005)	0.009** (0.004)
N	107,949	329,031	329,031
Cohorts	2007	2007–2009	2007–2009
Controls			
Middle school fixed effects		×	×
Cohort fixed effects		×	×
Individual characteristics			×

Notes: Each row reports the coefficients on predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column. The regression includes middle school and year fixed effects and is estimated on a subsample of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year and were wage earners in the DADS data 2016–2018 that excludes the students who attended Grade 9 in a *département* with a high contribution to the CAP channel. The coefficients are estimated based on administrative data from the French Ministries of Education and Higher Education and the Insee. Standard errors are clustered at the middle school \times year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

Table 1.E3 – Reform’s Reduced-Form Impact on Vocational Students’ Outcomes - Wage Earners Only, *Départements* without Class Size Adjustment

	Baseline mean	Reduced-form estimates	
	(1)	(2)	(3)
Matriculation in vocational track			
CAP	0.154	0.034** (0.015)	0.035** (0.014)
Seconde professionnelle	0.011	0.847*** (0.024)	0.848*** (0.024)
Class size	13.554	-0.539 (0.395)	-0.617 (0.393)
Educational outcomes			
Access to 12 th grade	0.407	0.138*** (0.021)	0.136*** (0.020)
High school graduation	0.319	0.084*** (0.016)	0.080*** (0.015)
Dropout	0.143	0.023** (0.010)	0.025*** (0.010)
Access to first year STS	0.093	0.022*** (0.007)	0.021*** (0.007)
Access to first year university	0.029	0.012*** (0.004)	0.011*** (0.004)
Access to second year STS	0.067	0.016*** (0.006)	0.015*** (0.006)
Access to second year university	0.003	0.002 (0.006)	0.002 (0.006)
STS graduation	0.049	0.005 (0.005)	0.004 (0.005)
Labour market outcomes			
Long-term contract	0.515	0.037*** (0.011)	0.037*** (0.011)
Log annual wage	9.361	0.027 (0.022)	0.030 (0.021)
Log hourly wage	2.496	0.005 (0.006)	0.006 (0.005)
N	86,502	263,617	263,617
Cohorts	2007	2007–2009	2007–2009
Controls			
Middle school fixed effects		×	×
Cohort fixed effects		×	×
Individual characteristics			×

Notes: Each row reports the coefficients on predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column. The regression includes middle school and year fixed effects and is estimated on a subsample of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year and were wage earners in the DADS data 2016–2018 that excludes the students who attended Grade 9 in a *département* with a high contribution to the class size channel. The coefficients are estimated based on administrative data from the French Ministries of Education and Higher Education and the Insee. Standard errors are clustered at the middle school \times year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

1.E.3 Reform Effects by Pre-Reform Student Type

Depending on their pre-reform type, students were differently impacted by the reform. For students who would have successfully completed a 2+2-year BEP in the absence of reform, the introduction of the *seconde professionnelle* programs solely reduces the time to graduation. To the contrary, for the students who would have undertaken a 2-year BEP only, or would have attempted to matriculate in Grade 11 or 12 and dropped out before completing the program, the treatment reduces the time to graduation but also affects their secondary schooling attainment. To better understand how this treatment heterogeneity may have impacted students' outcomes, I perform a prediction of the students' pre-reform type, and estimate Equation (1.1) for each predicted type separately.

Model specification and estimation. I assume that the students' pre-reform types are determined by the utility that they derive from being of the CAP, 2-year BEP or 2+2-year BEP type. The utility that student i derives from choosing alternative j is assumed to be a linear function of student i 's covariates X_i , which include origin middle school and Grade 9 fixed-effects, gender, age dummies, parental occupation dummies, and dummies indicating the student's twentile of prior achievement level:

$$U_{i,j} = \alpha_j X_i + \epsilon_{i,j} \quad \forall i, j \quad (\text{A.1})$$

I use a multinomial logistic specification to estimate model A.1. Assuming that the $\epsilon_{i,j}$ are type 1 extreme value distributed, the probability for a student to be of a given pre-reform type conditional on its covariates has a logistic form (McFadden et al., 1973):

$$P(\text{Type}_i = j | X_i) = \frac{\exp(\alpha_j X_i)}{\sum_{k=1}^J \exp(\alpha_k X_i)} \quad \forall i, j \quad (\text{A.2})$$

Using a maximum likelihood approach, I seek to find the set of α coefficients that maximize the log likelihood function $\ln(L(\alpha))$. To reduce dimensionality, I add to the objective function a penalty on the sum of the absolute values of the coefficients α , with regularization parameter λ :

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} (-\ln(L(\alpha)) + \lambda \sum_{x=1}^{X_n} |\alpha|) \quad (\text{A.3})$$

A training sample of pre-reform students that accounts for 30 percent of the pre-reform students is used to estimate the $\hat{\alpha}$ coefficients. The λ parameter is tuned based on a 10-fold cross validation procedure. I pick the highest λ value such that the error of prediction lies within one standard deviation of the minimal error (Hastie et al., 2009).

Eventually, using the $\hat{\alpha}$ coefficients, I can predict the pre-reform type for both pre- and post-reform students.

Goodness of fit. On average, 65.6 percent of the pre-reform students from the test sample are assigned their observed type. The model does a better job at predicting the CAP type (81.8 percent of predicted CAP students are indeed CAP students) and the 2-year BEP type (65.9 percent of predicted 2-year BEP students are indeed 2-year BEP students) than at predicting the completion of the 2+2-year BEP path (56.6 percent of predicted 2+2-year BEP students are indeed 2+2-year BEP students).

Table 1.E4 – Reform’s Reduced-Form Impact on Vocational Students’ Outcomes, by Predicted Pre-Reform Agent Type

	CAP		2-year BEP		2+2 year BEP	
	Baseline (1)	Estimate (2)	Baseline (3)	Estimate (4)	Baseline (5)	Estimate (6)
Matriculation in vocational track						
CAP	0.860	0.087*** (0.032)	0.102	0.030* (0.017)	0.081	0.040*** (0.013)
Seconde professionnelle	0.000	0.084*** (0.018)	0.012	0.892*** (0.029)	0.009	0.819*** (0.041)
Educational outcomes						
Number of years of secondary schooling	2.339	-0.072 (0.085)	3.012	-0.274*** (0.034)	3.484	-0.509*** (0.052)
Access to 12 th grade	0.062	0.015 (0.024)	0.377	0.173*** (0.023)	0.635	0.095*** (0.028)
High school graduation	0.031	0.000 (0.012)	0.274	0.099*** (0.016)	0.597	0.084*** (0.020)
Dropout	0.249	0.012 (0.046)	0.153	0.018* (0.009)	0.048	0.019 (0.013)
Access to first year STS	0.007	-0.003 (0.005)	0.077	0.035*** (0.008)	0.184	0.037** (0.015)
Access to first year university	0.003	-0.001 (0.004)	0.027	0.014*** (0.004)	0.040	0.010 (0.007)
Access to second year STS	0.004	0.005 (0.004)	0.052	0.025*** (0.005)	0.142	0.027** (0.013)
Access to second year university	0.000	-0.001 (0.004)	0.002	0.001 (0.005)	0.006	0.004 (0.013)
STS graduation	0.002	0.005* (0.003)	0.035	0.007 (0.005)	0.118	0.000 (0.012)
Labour market outcomes						
Long-term contract	0.421	-0.032 (0.026)	0.508	0.017** (0.009)	0.569	0.037 (0.026)
Log annual wage	8.987	0.023 (0.079)	9.387	0.012 (0.028)	9.595	0.042 (0.042)
Log hourly wage	2.437	0.001 (0.020)	2.497	0.011** (0.005)	2.514	-0.009 (0.009)
N	8,969	27,310	82,953	248,566	22,123	66,341
Cohorts	2007	2007– 2009	2007	2007– 2009	2007	2007– 2009
Controls						
Middle school fixed effects		×		×		×
Cohort fixed effects		×		×		×

Notes: Each row reports the coefficients on predicted exposure to the reform from the OLS estimation of Equation (1.1) for the different outcomes listed in the first column, performed for each pre-reform student type separately (see definition in Section 1.5.2). The regression includes middle school and year fixed effects and is estimated on the universe of students who were enrolled in Grade 9 between 2007 and 2009 and who attended a vocational track program the following year and were wage earners in the DADS data 2016–2018, based on administrative data from the French Ministries of Education and Higher Education and from the Insee. Standard errors are clustered at the middle school \times year level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

1.E.4 Exploring the Gender Heterogeneity of the Reform’s Labor Market Impact

Table 1.E5 – Male-Dominated vs. Female-Dominated Vocational Specialisations

	Male Students (%)	DNB Candidates (%)	DNB rank	<i>Bac</i> graduate (%)	First year STS (%)	First year university (%)	N Students
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Male-dominated specialisations							
Industrial engineering	97	85	27	29	10	2	46,213
Civil engineering	92	71	29	18	5	1	16,127
Industrial transformation	72	67	29	17	4	1	10,374
Agriculture and natural landscapes	72	48	25	0	1	0	502
Gender-balanced specialisations							
Management	45	84	27	34	10	4	43,056
Female-dominated specialisations							
Community services	29	78	24	9	1	1	3,324
Care	21	86	34	25	4	3	31,715
Communication	12	90	25	37	10	5	17,903
Soft materials	6	76	25	21	5	3	5,065

Notes: The table reports the share of male students, DNB candidates, the average DNB rank (for DNB candidates only) and the number of registered students by family of vocational specializations before the reform. The figures are computed on the universe of students who were enrolled in Grade 9 in 2007 and who attended a vocational track program the following year and were wage earners in the DADS data 2016–2018, based on administrative data from the French Ministries of Education and Higher Education.

Table 1.E6 – Wages and Employment Rates Nine Years After Grade 9: Male-Dominated vs. Female-Dominated Vocational Specializations

	Employed (%)			Annual Wage (euros)			Hourly Wage (euros)		
	<i>Baccalauréat</i> graduates			<i>Baccalauréat</i> graduates			<i>Baccalauréat</i> graduates		
	No	Yes	Diff (%)	No	Yes	Diff (%)	No	Yes	Diff (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Male-dominated specialisations									
Industrial engineering	73	84	15	17,031	21,314	25	13	14	6
Civil engineering	73	84	16	16,125	20,000	24	12	13	4
Industrial transformation	67	85	28	15,456	21,259	38	12	14	9
Agriculture and natural landscapes	61	100	63	13,673	10,430	−24	12	9	−18
Gender-balanced specialisations									
Management	65	80	22	14,121	17,059	21	12	12	2
Female-dominated specialisations									
Community services	61	71	16	13,443	15,767	17	12	12	1
Care	67	76	13	14,256	16,912	19	12	13	7
Communication	62	76	23	12,300	15,169	23	12	12	2
Soft materials	60	81	34	12,117	14,772	22	12	12	0

Notes: The table reports the share of employed students as well as the average annual and hourly wages of pre-reform students nine years after Grade 9, separately for each family of vocational specializations, and for *baccalauréat* graduates and non-graduates. The figures are computed on the universe of students who were enrolled in Grade 9 in 2007 and who attended a vocational track program the following year and were wage earners in the DADS data 2016, based on administrative data from the French Ministries of Education and Higher Education and the Insee.

Chapter 2

How Do Elite Post-Secondary Programs Select their Students? *

This paper uses data from the centralized system of college admissions in France to estimate selective programs' preferences over applicants. To assess how selective programs weigh each applicant's characteristic when making admissions decisions, we assume that they rank applicants based on a score value that depends linearly on their grades and demographic characteristics, among which gender, social background and geographical origin. Using a program-specific logistic lasso estimation, we find that selective programs give preferential treatment to geographically-close applicants. Applicants to selective programs located in their own high school or district are given an advantage that is equivalent to a 0.3 standard deviation increase in each of the applicant's grades, for the programs with the highest preference for geographically-close applicants. To assess how programs' preferences for applicants' demographic characteristics affect the diversity of their student body, we simulate the impact of a policy ban on demographic characteristics, and find that the ban only moderately affects the average composition of admitted students to selective programs, but can have substantial effects on the geographical composition of the student body for some types of programs.

Keywords: *higher education, stratification, elite programs, admission rules*

*This paper is jointly written with Gabrielle Fack, Julien Grenet and YingHua He. It greatly benefited from discussions and helpful comments from Élise Huillery and Arnaud Maurel.

Introduction

The fairness of access to elite post-secondary programs is a highly sensitive issue, as these programs are the main pathway to high level executive occupations, and therefore to dominant social positions. Despite the critical importance of understanding the determinants of students' access to these programs, little is known about elite post-secondary programs' preferences over students and the extent to which they contribute to the underrepresentation of certain groups of students in their student body.

Prior research has investigated how post-secondary programs competing for students strategically set their tuition fees (Andrews and Stange, 2019; Fu, 2014) or allocate their budget between consumption amenities and academic spending (Jacob et al., 2018) or course offerings (Thomas, 2019). Regarding admission policies, most studies have focused on affirmative action programs.¹ Beyond the context of affirmative action, the way in which selective programs weigh students' prior academic achievement and demographic characteristics, such as gender, socioeconomic background, and ethnic or geographical origin when making admission decisions has not been previously investigated in a systematic manner. As a consequence, little is known about programs' preferences for diversity among the students they admit, despite the central aspect of this question for policies seeking to increase diversity in higher education programs.

Recently, a large number of legislative and media debates have questioned the fairness and transparency of post-secondary programs' admission rules in general contexts that do not involve explicit affirmative action policies. The most prominent example is the *SFFA v. Harvard* case, in which the plaintiffs claimed, based on Asian American students' stronger academic records, that their lower admission rate compared to that of white students reflected race-based discrimination at entry into Harvard.

The scientific debate that surrounds this case has revealed the major empirical challenges that researchers face when trying to identify the part of observed social and ethnic stratification that cannot be imputed to the students' achievement level and skills. In the first place, the correlation between students' demographic characteristics (race in the context of the Harvard case) and their

¹In affirmative action settings, discussions have assessed the legality of giving preferential treatment to some groups of students. The most famous legal debate is relative to the legality of affirmative action in higher education in the US. In 1978, the Supreme Court ruled that programs are allowed to implement affirmative action policies to serve a diversity objective in the program, but not to correct for historical disadvantages of a minority.

achievement level prevents from concluding that differential admission chances across groups of students are a result of explicit discrimination. In addition, controlling for students' observed achievement may not be sufficient to bring evidence of a preferential treatment. As noted by Arcidiacono et al. (2020), programs may value unobservables that are correlated with students' demographic characteristics. For instance, in the event an applicant's high school of origin is unobserved, but social background is observed, programs may use social background as a proxy for high school quality. In this case, the valuation of students' demographic characteristics could reflect the contextualization of applicants' grades, i.e., a form of statistical discrimination, rather than taste-based discrimination. To isolate intrinsic preferences for a given group of students, Bhattacharya et al. (2017) showed that it is possible to rely on an approach that allows to identify taste-based discrimination under the assumption that students who have a better profile based on their observables also do better based on their unobservables.

In this paper, we show that preferential treatment for certain groups of students can be identified under weaker assumptions, when it is possible to exploit, in addition to all information that programs may use when making admissions decisions, data on students' achievement level at a national end of high school blind test. We use data from the centralized system of post-secondary admissions in France *Admissions Post-Bac* (APB), which assigns applicants to programs using a variant of the college-proposing deferred acceptance algorithm. For post-secondary admissions, non selective programs have to rank applicants based on exogenously defined rules, while selective programs have full discretion about their admission rules. Therefore, selective programs' preferences over applicants can be estimated based on their rankings of applicants. Although our approach could be extended to any type of selective post-secondary program, we focus on elite post-secondary programs (*Classes préparatoires aux grandes écoles*, hereafter CPGE), which are two-year programs that prepare their students to elite graduate schools, called *grandes écoles*. Unlike other post-secondary programs, they have a well-defined objective function, which is to maximize the share of their students who succeed at the entry exam to top-tier elite graduate colleges. Recently, the fairness and transparency of their admissions rules has been questioned, and the option of forbidding the use of high school identities and applicants' demographic information when making admissions decisions has been considered. We shed some light on these programs' admission rules by estimating their preferences over applicants, and providing evidence that their valuation of applicants' demographic characteristic does not solely reflect the contextualization of their grades. To assess the extent to

which their preferences affect the diversity of their student body, we simulate the impact of a policy ban on applicants' demographic characteristics, and show that this policy ban would not affect the quality of prediction of their applicants' probability of success at the entry exam to elite graduate colleges.

This paper makes several contributions to the literature on post-secondary admissions. Our first contribution is to propose a method to estimate the relative valuation of each aspect of the applicants' profiles by the programs from the programs' rankings of applicants. In the French centralized post-secondary admissions system, unlike in most decentralized admission schemes, programs' rankings of applicants are observed. In addition, since elite programs have discretion about their admission rules, their rankings of applicants do not result from deterministic priority rules such as those used in most centralized school choice systems. However, an important feature of the French system is that it allows researchers to access the same set of information on applicants as that used by programs when generating their rankings of applicants. We exploit this unique setting to reverse engineer elite programs' rankings of applicants. Our model of programs' rankings of applicants builds on the decision-making tool (DMT) made available to the programs on the APB platform. The DMT can produce automated rankings of applicants based on a linearly additive score computation that relies on a set of coefficients on grades and on applicants' demographic characteristics submitted by the programs. Following this approach, we assume that elite programs rank applicants based on a score value that can be linearly decomposed between a grades-based score (a weighted average of the applicants' grades) and a demographic score (a sum of positive or negative weights on applicants' gender, socio-economic background, high school and region of origin). Using a program-specific logistic lasso estimation on the universe of elite programs that participated to the 2016 centralized post-secondary assignment procedure, we estimate how these programs weigh applicants' demographic characteristics while controlling for all the information on applicants' achievement level available to the programs.

We find that elite programs give preferential treatment to geographically-close applicants. Applying to an elite program located in one's own high school or district yields a score increment that is equivalent to a 0.3 standard deviation increase in the grades-based component of the score, for the elite programs with the highest valuation of geographically-close applicants. We provide evidence that the elite programs with the strongest valuation of geographically-close applicants tend to be highly selective science programs located in Paris.

We check that the preferential treatment given to geographically-close applicants reflects a genuine preference for this group of applicants rather than the contextualization of their grades. To do so, our approach consists in using the applicants' grades at the end of the French high school national examination (*baccalauréat*) as a directly comparable measure of the applicants' achievement levels. As programs do not have access to this information when making admission decisions, we mimic their behavior by predicting the *baccalauréat* grades using applicants' continuous assessment grades and school of origin. We show that the coefficients obtained on applicants' demographic characteristics when predicting their ranking are robust to the inclusion of these controls. This suggests that weights on applicants' demographic characteristics do not account for unequal predicted performance levels across groups of students. However, the positive weights on geographically-close applicants would be consistent with risk-coping strategies from the CPGE programs, which could possibly favor applicants from their high school based on inside information obtained through informal interactions with their Grade 12 teachers.

Our second contribution is to assess the impact of a policy ban on applicants' demographic characteristics. We run counterfactual simulations of how programs would rank applicants in a limited information setting where the only available information would be on applicants' high school grades. To produce the counterfactual rankings, we let elite programs try and replicate their observed ranking of applicants using high school grades only, which assumes that programs have perfect information on the correlation between high school grades and the demographic characteristics that they value, and therefore provides the lower bound effect of the ban on elite programs rankings. To assess the impact of this ban on stratification among these programs, we replicate the APB centralized assignment mechanism based on the counterfactual rankings of applicants and generate the applicant-program match that would be obtained under the assumption that students do not modify their application behavior in reaction to the change in programs' admissions rules. We find that the policy ban would permit a modest, yet significant increase in terms of diversity of students' high school and district of origin. In particular, the ban would achieve a decrease from 9.1 percent to 8.4 percent in the share of applicants admitted to an elite program located in their origin high school. We give suggestive evidence that this policy ban would not be costly to elite programs, under the assumption that their objective function involves maximizing the share of their students accessing a top-tier graduate school. We do so by showing that the prediction accuracy of access to top-tier elite graduate schools based on the scores assigned

to applicants is not lower in the counterfactual scenario compared to the baseline scenario.

Related Literature. This paper is related to several strands of the literature. First, it contributes to the literature on post-secondary admissions policies. In centralized systems, a significant variety of institutional rules have been devised to determine how to assign applicants to programs. On top of the mechanism used, one important feature of an assignment system is the degree of discretion left to programs when generating the rankings of applicants. They may have to admit applicants based on exogenously defined formulas (e.g. the applicants' GPA scores in Scandinavian countries or an end of high school exam-based score in Tunisia (Luflade, 2017)). On the contrary, programs may have full discretion on the rankings of applicants, in which case their admission policies can have a direct effect on the diversity of their student body.

In countries with explicit race- or caste-based affirmative action policies and where programs have discretion upon students' rankings, namely Brazil, India and the US, prior research on admission policies has sought to find evidence on the scope of affirmative action implemented in post-secondary admissions, and produced empirical evidence showing that the preference for diversity is particularly marked in top-tier programs (Kane, 1998; Arcidiacono, 2005; Howell, 2010; Arcidiacono et al., 2015; Arcidiacono and Lovenheim, 2016; Long, 2004; Epple et al., 2008). Changes in the legislation regarding affirmative action have allowed to assess the extent to which programs manage to achieve their diversity objective under bans on students' demographic characteristics (Chan and Eyster, 2003; Long and Tienda, 2008). Closely related to our paper, Antonovics and Backes (2014) assess how the undergraduate programs of the University of California reacted to a ban on race-based affirmative action. The authors find evidence that the ban led to substantial adjustments in the programs' weighting of applicants' characteristics, with increases in the relative weight given to GPA and parental background compared to SAT scores. In contrast to this paper, we assess the impact of a ban on applicants' demographic characteristics using a structural approach and in a more general context without explicit affirmative action policies.² Such contexts have been much less studied, with the notable exception of the *SFFA v. Harvard* case studied by Arcidiacono et al. (2020).

Similarly to race-based preferences, the investigation of the extent of gender-based preferential treatment in higher education admissions has been mostly confined to explicit affirmative settings

²As data collection on race is prohibited in France, we do not include this dimension in our analysis, but rather focus on gender, SES and school and district of origin.

(see for instance Bagde et al. (2016) in India). Aside explicit affirmative action settings, Breda and Ly (2015) and Breda and Hillion (2016) find evidence of a preference for gender balance in competitive examinations at the graduate level in France. Their results suggest that female (male) candidates receive preferential treatments in examinations in a male-dominated (female-dominated) subject—a result that is consistent with this paper’s findings at the post-secondary level.

Compared to race and gender, little is known about the role played by students’ geographical background, including their place of residence and high school of origin, in the admission decisions of higher education programs.³ In the US, there is evidence that state universities tend to favor in-state students in response to state pressure (Groen and White, 2004), and that declining state appropriations encourage them to grow out-of-state students’ enrollment, due to the higher tuition fees that they can charge to these students (Jaquette and Curs, 2015). In this paper, we find a similar preference of elite post-secondary programs for geographically-close students, which cannot however be attributed to these programs’ funding strategies, since they are fully publicly funded and do not charge tuition fees to their students.

Lastly, this paper relates to the literature on preference estimation in centralized school choice and college admissions. Prior research has essentially focused on the estimation of students’ preferences (see, for instance, Avery et al. (2013) for the estimation of students’ preferences over colleges in a decentralized system). In centralized systems, the estimation of students’ preferences has been facilitated by the availability of applicants’ rank-order lists of colleges and by the properties of the assignment systems. Assuming that applicants’ rank-order lists of programs are truthful reports of their preferences, part of the literature has estimated students’ preferences for schools (Abdulkadiroğlu et al., 2020) or undergraduate programs (Drewes and Michael, 2006; Kirkebøen, 2012)). Departing from the truth-telling assumption, which could fail in settings where there is a restriction on the number of choices that the applicants can submit (Haeringer and Klijn, 2009; Calsamiglia et al., 2010) two alternative strands of the literature have used the stability of the assignment mechanism (Akyol and Krishna, 2017; Fack et al., 2019; Bucarey, 2018) or an approach formalizing the submitted rank order list as the one maximizing each applicant’s expected utility – which assumes that applicants understand the assignment mechanism and the environment that

³Whether and how programs should evaluate students’ prior educational context, namely their high school of origin and the geographical area where they grew up, is however a highly contentious question, as this form of statistical discrimination may harm the admissions chances of high achievers from low quality high schools and disadvantaged neighborhoods. While in Quebec, a 1996 ruling nationally unified the way to contextualize the high school GPA of applicants, the Canadian University of British Columbia has been forbidden to use high school identities in admissions for undergraduate programs in economics.

they face (Agarwal and Somaini, 2018) – to recover students’ preferences for high schools or colleges. We complement this literature by focusing on the programs’ side. Compared to the estimation of students’ preferences, the estimation of programs’ preferences presents fewer methodological challenges, as programs do not face incentives to behave in a strategic manner when ranking applicants, so that truthful preference revelation is a credible assumption.

This paper is organized as follows. Section 2.1 provides institutional background on the French higher education system and describes the main features of the student assignment to post-secondary programs. Section 2.2 presents the data sources that we use, along with summary statistics on applicants to CPGE programs and the rankings of applicants produced by CPGE programs. Section 2.3 presents the model used for the estimation of programs’ preferences. Section 2.4 describes how programs weigh students’ academic achievement level and demographic characteristics. In Section 2.5, we simulate the impact of a policy ban on applicants’ demographic characteristics on programs’ composition and applicants’ outcomes. We discuss our results in Section 2.6. Section 2.7 concludes.

2.1 Institutional Background

This section provides institutional background on the French education system and describes the main features of the student assignment to post-secondary programs.

Structure of the French education system. In France, education is compulsory from the ages of 6 to 16 years. At the end of primary school, students start their secondary schooling in a common path taught in middle schools (Grades 6 to 9). After Grade 9, they join either a vocational high school, or a general and technological high school, based on their achievement level and preferences. The general and technological high school paths last three years (Grades 10 to 12). In Grade 11, general and technological high school students are tracked between the the humanities, economics, science, and technological tracks, and further specialize when entering Grade 12, by choosing an elective course.⁴

At the end of high school, students take the national *baccalauréat* examinations, which are

⁴Students in science high school track can opt for an earth and life science, math, physics, computer science or environment elective. Students enrolled in the economics track are offered electives in math, social and political science or advanced economics. Humanities students can select an advanced foreign language, law, arts, or math elective.

specific to each high school path and further vary with the students' track and elective courses. The examination process spans over the last two years of high school. Students are tested on the minor academic disciplines of their track at the end of Grade 11, with the so-called preliminary *baccalauréat* examinations (*épreuves anticipées*).⁵ Their achievement in the core subjects of their track is assessed with the final examinations at the end of Grade 12. Students who pass the general, technological or vocational *baccalauréat* are awarded the corresponding *baccalauréat* degree, and are automatically eligible for higher education.

Post-secondary admissions in France. Until 2016, the post-secondary admissions procedure was managed through the *Admissions Post-Bac* (APB) online platform for most selective and non selective post-secondary programs.⁶ The system was established in 2008 by the French Ministry of National Education, Higher Education and Research (MENESR) to simplify the admission process by gathering together nearly all post-secondary programs. It collects applicants' rank-ordered list of choices and programs' priority ranking of applicants, and, with this information on both sides, it sequentially assigns students to programs using a variant of the college proposing deferred acceptance algorithm (Haeringer and Iehlé, 2019). Appendix 2.A details the specificities of the APB matching mechanism.

Each year, over 750,000 high school students connect to the APB website to gather information on post-secondary programs, submit their rank-ordered list of choices, and participate in the admission phases. The programs to which students can apply fall into two broad categories, each accounting for approximately half of first-year post-secondary enrollment: the selective and the non selective programs. Non selective university programs (*Licence*) are open to all students who hold the *baccalauréat*. They cannot select students based on their academic background.⁷ Selective programs, including *Classes préparatoires aux grandes écoles* (CPGE), two-year technical programs and post-secondary *grandes écoles* are free to select their students based on their own criteria, including students' past academic records. Selective programs are required to return to

⁵Preliminary *baccalauréat* examinations cover oral and written French for all students, as well as history and geography for science track students, earth and life science for economics track students, and science for humanities track students.

⁶Since the 2018 reform of admission into higher education, the Parcoursup procedure has replaced the APB procedure. Our study is based on the college admission data for 2016, i.e., prior to this reform.

⁷When the number of applications exceeds the number of seats available, students from the university's academic area (*académie*) are given priority over those from elsewhere and, within the group of local applicants, students who have ranked the program at a higher position in their list are given priority; remaining ties are broken using a random lottery.

the centralized clearinghouse a ranking of all or part of the students who have applied to them, without knowing at which position they were ranked by applicants.

The APB website is open from December 1 to mid-September of the following year. The centralized admission procedure comprises three main phases:

1. Registration Phase (January 20 – May 31): students are requested to register on the website and to select the programs they wish to apply to before April 2. By the end of May, they must rank and confirm their applications (at most 12 choices per program type and at most 36 in total).
2. Admission Phase (June 8 – July 14): selective programs examine the applications and return their priority ranking of applicants to the centralized clearinghouse. There are three rounds between early June and mid-July. In each round, the matching algorithm is executed; students are offered a seat at the highest-ranked choice at which they received a positive answer; lower-ranked choices are automatically canceled while waiting lists are created for students who have a chance of being accepted if seats are freed up in the next rounds. In the first two rounds, each student can either (i) accept the offer and withdraw her other applications; (ii) accept the offer but wait until the next round before withdrawing her other applications; (iii) turn down the offer but wait until the next round before withdrawing her other applications; or (iv) withdraw all her applications. In the third round, students are asked to definitively accept or reject their offer, and to finalize their enrollment.
3. Supplementary Phase (*procédure complémentaire*): there exists an additional procedure for students who remain unassigned after the admission phase, and for those who did not register in time for the admission phase. This phase runs during the summer until the end of September.

APB decision-making tool. To assist selective programs in their ranking of applicants, a decision-making tool (DMT) is made available to the programs on the APB platform. The DMT can produce automated rankings of applicants based on a set of coefficients on grades and on applicants' demographic characteristics submitted by the programs. To produce the rankings of applicants, the DMT assigns to each applicant a linearly additive score that is computed as a weighted average of the applicants' grades, and potentially adjusted with negative or positive weights on other dimensions of the applicants' profile, such as their high school of origin or whether they have repeated a grade. Programs may then submit this raw ranking of applicants on the

platform, or a revised version of the raw ranking, or produce an entirely different ranking of applicants.

CPGE programs. The CPGE programs are the most selective post-secondary programs. These two-year programs are located in high schools and specialize in science, business or humanities. They gather eight percent of post-secondary students, and their primary purpose is to prepare students to the entry exam to elite graduate schools (the *grandes écoles*, who gather 25 percent of graduate students). Although most elite graduate schools recruit students outside the CPGE pool of students, the CPGE are the main pathway to the most selective elite graduate schools. In 2016, 38 percent of elite graduate schools students had undertaken a CPGE program at the post-secondary level, and this was the case of 69 percent of the students in the 10 percent most selective elite graduate schools, hereafter called top-tier graduate schools.⁸

The CPGE do not charge any tuition fees to their students, who however have to administratively enroll in a partner university, which charges fees of 170 euros per year, and to pay a yearly national student contribution of 92 euros, the *Contribution de vie étudiante et de campus*.

2.2 Data and Summary Statistics

To estimate CPGE programs' preferences and assess how alternative admission criteria would impact the social, gender and spatial composition of their student body, we use information from the APB centralized admission platform. The data cover all applicants and programs that participated in the APB centralized admission procedure in 2016.

2.2.1 Data Sources

Applicants' characteristics We estimate programs' preferences using all the information on applicants that programs have at their disposal when producing the rankings of applicants. The access to the comprehensive data from the APB platform was provided to us by the statistical office of the French Ministry of Higher Education (MESRI-SIES). Along with basic socio-demographic information including gender, district of origin, SES, and eligibility to a need-based scholarship, the data contain detailed information on applicants' academic profile. In particular, they provide

⁸All figures mentioned in this paragraph can be found in Bonneau et al. (2021). Selectivity is computed based on the students' average percentile rank at the general *baccalauréat* examinations.

information on the students' secondary schooling history, including their high school of origin, high school track and elective subjects chosen, whether they repeated a grade, the grades obtained at the preliminary *baccalauréat* examinations, as well as the continuous assessment grades in all subjects taken in Grade 11 and Grade 12 (one grade per subject and quarter).

***Baccalauréat* performance data.** We enrich the APB data with information on applicants' grades on the *baccalauréat* exams, which are retrieved from the OCEAN database, managed by the statistical office of the French Ministry of Education (MENJS-DEPP). These data record the grades obtained by the *baccalauréat* candidates, for each subject. As the *baccalauréat* exam is anonymously and externally graded, the grades obtained in each subject provide a comparable measure of students' academic performance, in contrast with the continuous assessment grades that are not directly comparable across classes and high schools. It is worth noting that because the *baccalauréat* final exams take place in June, the information on students' *baccalauréat* grades are not available to the programs when they rank applicants, except for the grades obtained at the preliminary *baccalauréat* examinations.

Enrollment in higher education. APB applicants' enrollment status in higher education for the years 2018 and 2019, i.e., two to three years after the 2016 round of applications, is derived from the SISE administrative datasets of the French Ministry of Higher Education. This information is used to identify the students who access the top-tier *grandes écoles* at the graduate level.

Programs' characteristics. The APB data contain detailed information on each participating program, including the program type (university, CPGE, etc...), field of specialization (which, for CPGE programs, are humanities, science, and business) and location. The APB data further indicates for each selective program whether it used the DMT to generate its ranking of applicants.

Replication of the APB Matching. To perform counterfactual student assignment simulations, we use all the information necessary to replicate the APB student-program matching, i.e., (i) applicants' rank-order lists of programs, (ii) programs' rankings of students, (iii) programs' capacities. In addition, the data keeps track of all the offers received by each student, and whether the offer was accepted or not.

2.2.2 Student Characteristics

To assess how CPGE programs take into account students' demographic characteristics when making admissions decisions, we define synthetic indices of the students' socio-economic background and geographical origin.

Students' socio-economic background. Students' socioeconomic status (SES) is measured using the French Ministry of Education's official classification, which uses the occupation of the child's legal guardian to define four groups of SES: high (company managers, executives, liberal professions, engineers, intellectual occupations, arts professions), medium-high (technicians and associate professionals), medium-low (farmers, craft and trades workers, service and sales workers), and low (manual workers and persons without employment).

This first measure seeks to capture differences in the human capital associated to each occupation. In addition, we adopt a complementary measure of socioeconomic background that focuses on parental income. This second measure classifies as low-income all students that are eligible to a need-based scholarship in secondary education.

Students' geographical origin. To uncover spatial inequalities in the access to selective post-secondary programs, we compare applicants living in one of the five cities with the highest number of applicants to the CPGE programs (hereafter top five cities) to all other applicants to CPGE programs. We also compare the applicants applying in a program located in their high school of origin to the other applicants, and those applying in a program located in their district (*département*) of origin to the other applicants.

2.2.3 Sample Restrictions and Summary Statistics

Our sample consists of all students who applied to at least one CPGE program in the APB procedure for 2016.⁹ Out of the 761,855 students who participated in APB that year, 97,011, i.e., 12.7 percent, applied to at least one CPGE program. Our sample of analysis is further restricted to the 87,629 CPGE applicants for whom the high school track information is available, as this piece of information is required to implement our programs' preferences estimation approach (see Section 2.3 for more details).

⁹Data on non-CPGE applicants is also used to replicate the APB matching.

Characteristics of CPGE applicants. Table 2.1 presents summary statistics on the CPGE applicants as a whole. In 2016, 47 percent are females, and 9 percent come from the top five cities. Only 11 percent of the applicants are low-income students, and 5 percent are from a low SES background. Their academic background is very homogeneous: over 90 percent of the CPGE applicants obtained the general *baccalauréat*, and two thirds completed the science track in high school. Both high SES and top 5 cities applicants tend to exhibit higher *baccalauréat* grades and to come from a higher quality high school (as measured by the high school students' average *baccalauréat* rank) compared to their peers.

Tables 2.B1 to 2.B3 in Appendix 2.B further contrast applicants' characteristics across the three main types of CPGE programs. The main dimension of heterogeneity is the gender composition of CPGE programs: female students are less likely to apply to a CPGE specialized in a male-dominated field (they represent only 36 percent of science CPGE applicants), and are more likely to apply to a CPGE specialized in a female-dominated field (74 percent of applicants to humanities CPGE are female students).

Characteristics of CPGE programs. Table 2.2 provides information on the CPGE programs, and compares them to the other post-secondary programs that participated in the APB platform in 2016. In 2016, the data cover a total of 836 CPGE programs, accounting for 6.8 percent of the programs participating in the APB assignment procedure. Half of the CPGE programs are specialized in science (438 programs), while the other half is made of 267 business CPGE and 131 humanities CPGE.

Each CPGE has around 58 seats on average, and the vast majority manage to fill all their seats at the end of the assignment procedure (this is the case of more than 70 percent of the CPGE, compared to less than 50 percent for other programs). This implies that the CPGE rankings of applicants directly influence the composition of their student body. On average, each CPGE program receives 501 applications, and ranks 297 of them. There are therefore two stages at which they can contribute to the stratification of their student body: (i) when they decide *whether* to rank an applicant or not, and (ii) when they decide *how* to rank the applicant.

Descriptive evidence on ranked vs. unranked applicants. Table 2.3 contrasts the characteristics of ranked and unranked applicants to CPGE programs. Compared to unranked applicants, ranked applicants correspond more often to students applying in their district of origin and having

a higher academic achievement level, for all types of CPGE. Surprisingly, the average quality of the high school of origin is lower among ranked applicants compared to unranked ones.¹⁰

Descriptive evidence on high- vs. low-ranked applicants. Table 2.4 compares the characteristics of the applicants in each quartile of CPGE programs' academic rankings. For consistency with the literature's traditional ability rankings of students, we define the quartiles as a decreasing function of student rank, so that the fourth quartile corresponds to the highest ranked applicants, who tend to be higher achievers.

Better ranked applicants achieved a higher performance on the *baccalauréat* examinations. All types of CPGE also tend to rank better female students and students who originate from the program's district, and to rank lower low-income and low-SES students. Compared to other CPGE types, the science CPGE appear to be particularly keen to giving a premium to students from the high school where the CPGE is located.

2.3 How Each Program Ranks Its Applicants

To investigate how a CPGE program ranks its applicants, we develop a model that uses the information in the applicants' files. The model is program-specific and gives a score to each applicant based the applicant's grades and characteristics.

More specifically, similar to the DMT on the APB platform, to compute a score for an applicant, a program first calculates a weighted average of the applicant's grades in a subset of subjects. Its applicants may come from different high school tracks such as science, humanities, and economics. We therefore allow the program to have track-specific weights on the applicant's grades. In addition, the program may give positive or negative weights on applicant characteristics such as gender, high school, and SES.

2.3.1 A Conditional Logit Model

Let S_{ji} be the score that program j gives to applicant i , such that program j ranks applicant i higher than applicant i' if and only if $S_{ji} > S_{ji'}$. As we will have a scoring model for each program, we will simplify S_{ji} to S_i by omitting the subscript j .

¹⁰High school quality is measured using the average percentile rank at the *baccalauréat*, computed for general track students only and separately for each high school track (science, humanities, economics).

Let T be the total number of high school tracks, C be the number of applicant characteristics, and K be the maximal number of subjects that an applicant could have taken in high school. Note that we treat math in the science track as a subject different from math in the humanities track. Let x_{ik} be student i 's grade in subject k . Naturally, a student does not take all K subjects. We define $e_{ik} = 0$ and $x_{ik} = 0$ if i has never taken subject k ; otherwise, $e_{ik} = 1$ and $x_{ik} \geq 0$.

Applicant i 's score is determined as follows:

$$S_i = \sum_{t=1}^T \mathbf{1}(T_i = t) \left[\sum_{k=1}^K \beta_{tk} \frac{\sum_{k'=1}^K \beta_{tk'}}{\sum_{k'=1}^K \beta_{tk'} \mathbf{1}(e_{ik'} = 1)} x_{ik} + \sum_{c=1}^C \gamma_{tc} d_{ic} \right] + \epsilon_i \quad (2.1)$$

such that

$$\begin{aligned} \beta_{tk} &\geq 0 \quad \forall t, k \\ \sum_{k=1}^K \beta_{tk} &> 0 \quad \forall t \end{aligned}$$

where T_i is the high school track of i ; $\mathbf{1}(\cdot)$ is an indicator function; $\beta_{tk} \frac{\sum_{k'=1}^K \beta_{tk'}}{\sum_{k'=1}^K \beta_{tk'} \mathbf{1}(e_{ik'} = 1)}$ is the non-negative weight on the grade in subject k for students from track t ; γ_{tc} is the weight on characteristic c for students from track t , which can be positive, negative, or zero; d_{ic} is a dummy indicating whether student i is of type c ; and ϵ_i is a type I extreme value.

A few clarifications regarding equation (2.1) are in order. First, all weights are track-specific, which allows the program to treat students from different tracks completely differently. Second, the weight on grade x_{ik} , $\beta_{tk} \frac{\sum_{k'=1}^K \beta_{tk'}}{\sum_{k'=1}^K \beta_{tk'} \mathbf{1}(e_{ik'} = 1)}$, takes into account the subjects that student i has taken ($e_{ik'} = 1$). If the student has taken all subjects, the weight becomes simply β_{tk} ; otherwise, the weight on x_{ik} is greater than β_{tk} , because $\frac{\sum_{k'=1}^K \beta_{tk'}}{\sum_{k'=1}^K \beta_{tk'} \mathbf{1}(e_{ik'} = 1)} > 1$. In effect, the weights on subjects that have not been taken ($e_{ik'} = 0$) by the student are proportionally distributed to subjects that have been taken ($e_{ik'} = 1$). Moreover, it holds constant the total sum of all weights on the subjects that have been taken by i , which is equal to $\sum_{k'=1}^K \beta_{tk'}$. Third, the weights on student characteristics can be of any sign, reflecting that the program can value a type of students either positively or negatively.

2.3.2 Penalized Conditional Logistic Regression

Without loss of generality, we index students by their position in the program's ranking, i.e., $i = 1$ is the student ranked first, $i = 2$ is the one ranked second, and so on. Let $V_i = \sum_{t=1}^T \mathbf{1}(T_i =$

t) $\left[\sum_{k=1}^K \beta_{tk} \frac{\sum_{k'=1}^K \beta_{tk'}}{\sum_{k'=1}^K \beta_{tk'} \mathbb{1}(e_{ik'}=1)} x_{ik} + \sum_{c=1}^C \gamma_{tc} d_{ic} \right]$. The log likelihood of observing the realized ranking of students, $i = 1, \dots, N$, is therefore:

$$\ln L(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \ln \left[\prod_{i=1}^N \frac{\exp(V_i)}{\sum_{i' \geq i} \exp(V_{i'})} \right] = \sum_{i=1}^N \left(V_i - \ln \left[\sum_{i' \geq i} \exp(V_{i'}) \right] \right) \quad (2.2)$$

In the data, there are 103 different subject names, which becomes 515 unique subjects when we consider the high school track of each subject. To avoid overfitting, we adopt a LASSO approach for the logit model as follows:

$$\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}} = \arg \min_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \left(-\ln L(\boldsymbol{\beta}, \boldsymbol{\gamma}) + \lambda \sum_{t=1}^T \left(\sum_{k=1}^K |\beta_{tk}| + \sum_{c=1}^C |\gamma_{tc}| \right) \right)$$

Appendix 2.C details how we impose the non-negativity constraints on the parameters β_k .

To solve the constrained minimization problem, we employ the coordinate descent algorithm proposed by Friedman et al. (2010) and adapted to conditional logistic models by Reid and Tibshirani (2014). More details are provided in Appendix 2.C.

2.4 Programs' Preferences and Rankings of Students

This section presents the results from the estimation of the model described by Equation (2.1) based on the CPGE rankings of applicants. In Section 2.4.1, we discuss the main insights that can be derived from this approach about CPGE programs' preferences over applicants. In Section 2.4.2, we assess the extent to which the observed rankings of applicants by CPGE programs can be replicated using our preference estimates. In Section 2.4.3, we show that the weights on applicants' demographic characteristics do not reflect the contextualization of their grades.

2.4.1 Programs' Preferences

We estimate the weights that CPGE programs assign to applicants' grades and demographic characteristics using the program-specific penalized logistic lasso model described in Section 2.3. Our specification includes the following regressors: (i) continuous assessment grades in Grades 11 and 12, including, among others, math, physics, biology, engineering science, philosophy, foreign and ancient languages, history and geography, French, literature, economics (ii) preliminary *baccalauréat*

grades, including written and oral French, history and geography and science subjects (iii) applicants' demographic characteristics, including gender, eligibility to a need-based scholarship, SES, and (iv) a set of student \times program-specific dummies indicating whether the program is located in the student's high school, district, and academic region of origin.¹¹

Model fit. We measure the model's goodness of fit for program j by averaging the predicted probabilities that each pair of program j applicants is ranked in the observed order:

$$\frac{1}{\frac{n_j(n_j-1)}{2}} \sum_{i,i'>i} \frac{\exp(\hat{V}_i)}{\exp(\hat{V}_i) + \exp(\hat{V}_{i'})}$$

where n_j is the number of ranked applicants in program j and applicants receive ranks $i \in [1, n_j]$, where the preferred applicant is ranked one and the least preferred ranked applicant is ranked n_j .

This measure can take values between 0 and 1 and is independent of the size of the ranking. Our model achieves a mean goodness of fit across CPGE programs of 0.80, with 95 percent of the programs' goodness of fit ranging between 0.60 and 0.97.¹²

Applicants' estimated scores: academic performance. Figures 2.1 to 2.3 present the distributions across programs of the estimated weights on Grade 12 and preliminary *baccalauréat* grades, for science, business and humanities CPGE separately, and for the subjects receiving the highest weights only (see Table 2.D1 in Appendix 2.D for the weights on all subjects). As expected, we find that the different types of CPGE programs value the applicants' performance in the core subjects of the program's field of specialization. Science CPGE tend to put high weights on grades in math and physics, with a median weight of respectively 0.30 and 0.20, and, to a lesser extent, to humanities related subjects such as foreign language, philosophy and the results obtained at the French preliminary *baccalauréat*. Business CPGE tend to use frequently the Grade 12 math grade, but also the grades obtained in social sciences subjects, such as economics, philosophy, foreign language and history and geography. Humanities CPGE have more heterogeneous valuations of the different subjects: the median weight assigned to each subject never exceeds 0.11. Unsurprisingly,

¹¹Each grade is normalized so as to have mean 0 and standard deviation 1 in the population of applicants from the same high school track, and is indexed by the applicants' high school track.

¹²In Appendix 2.E, we show that the scoring approach described in Section 2.3 yields predicted rankings of applicants that are almost perfectly correlated with the observed rankings. The correlation coefficients are only marginally lower for the CPGE programs that do not use the DMT compared to those who use it, which we interpret as evidence that the CPGE programs who do not explicitly use the DMT however make use of similar ranking rules.

these programs take into account mostly humanities-related subjects (philosophy, history and geography, foreign language, and French preliminary *baccalauréat*).

Applicants' estimated scores: demographic characteristics. Figures 2.4 to 2.6 report the distributions of the estimated coefficients on applicants' characteristics, and Panel (a) of Table 2.5 reports the corresponding bottom and top quartile values. These coefficients can be interpreted as bonuses/malus on the grades-based score. As all grades are normalized to have mean 0 and standard deviation 1 in the sample of CPGE applicants from the same high school track, a bonus/malus of one on a given demographic characteristic can be interpreted as the equivalent of a one standard deviation increase/decrease in each of the grades obtained in the subjects used by the program in the grades-based score computation.

We find that less than 50 percent of the CPGE programs apply positive or negative weights to applicants on the basis of their demographic characteristics.¹³ For gender and socio-economic background, the magnitude of the corresponding coefficients is low, and rarely exceeds 0.3 standard deviations in absolute value. Humanities CPGE, however, tend to be more inclined to value applicants' demographic characteristics (up to 50 percent of the programs) compared to business and science CPGE (less than 40 percent of the programs). For all types of CPGE, the share of programs that place a positive weight on students' low-income status is slightly higher than the share of programs penalizing them. In business and humanities CPGE, this however coincides with a higher propensity to assign negative weights to low-SES applicants. The valuation of female applicants by CPGE programs is very heterogeneous across CPGE types. While the share of business CPGE programs assigning a positive weight to female students (39 percent) is higher than the share giving them a negative weight (26 percent), the reverse pattern is observed in humanities CPGE, with only 40 percent of the CPGE assigning them a positive weight while 53 percent assign them a negative weight.

All types of CPGE tend to have a strong preference for applicants that originate from the same school and same district as the program, and/or top five cities applicants. This is particularly marked for science and humanities CPGE programs, half of which put a strictly positive coefficient on students applying in their origin high school (vs. approximately 45 percent for business CPGE). Unlike the weights applied to other applicants' characteristics, those assigned to applicants from same

¹³More specifically, less than 50 percent of CPGE programs apply weights that are higher, in absolute value, than one percent of a standard-deviation of the grades-based score to their applicants' demographic characteristics.

high school and same district as the program can be very large. The top quartile weight assigned to applicants from the high school ranges from 0.16 (0.18) standard deviations for humanities (business) CPGE to 0.30 standard deviations for science CPGE. Similarly, district applicants can receive large positive weights, with a top quartile weight of 0.07 standard deviations in humanities CPGE, 0.10 standard deviations in business CPGE and 0.16 in science CPGE. Such behavior does not necessarily correspond to an inner preference for geographical proximity. As discussed earlier, CPGE programs might try to cope with the risk inherent to the selection of applicants on a limited set of observables by favoring applicants from their high school, for which they can access inside information, for instance through informal interactions with their Grade 12 teachers.

Characterizing the programs with high valuation of applicants' demographic characteristics. To characterize the type of CPGE programs that tends to take into account their applicants' demographic characteristics when computing their scores, we compare, for each demographic characteristic of interest, the CPGE programs assigning a weight that exceeds (resp. lies below) the top (resp. bottom) quartile weight to all other CPGE programs. Tables 2.6 and 2.7 report the OLS coefficients of the regression of a dummy indicating whether the CPGE has a highly negative (resp. positive) valuation of the considered demographic characteristic on the CPGE characteristics.¹⁴ Our results suggest that the CPGE programs with the strongest valuation of geographically-close applicants tend to be highly selective science CPGE located in Paris. Compared to the baseline science CPGE located outside Paris and belonging to the first quartile of selectivity, the CPGE programs located in Paris are more than 23 percentage points more likely to assign a top quartile weight to applicants from the same high school, district, or academic district. By contrast, business and humanities CPGE programs are 6.7 to 20.4 percentage points less likely to assign a top quartile weight to each of these characteristics, compared to the baseline CPGE programs.

2.4.2 Programs' Rankings of Students

Using the set of estimated weights on grades and applicants' demographic characteristics, we can recover each program's ranking of applicants by computing the scores as described in Section 2.3, and assigning a rank of one to the highest scoring applicant, a rank of two to the second highest scoring applicant, and so on and so forth. To account for the random component of each program's

¹⁴As bottom quartile weights are very close to zero for all of the demographic characteristics of interest, we comment on the results relative to CPGE assigning top quartile weights only.

scoring function, we randomly draw, for each program j , 300 vectors of type-1 errors u and predict the applicant i 's score in simulation s as follows:

$$\hat{V}_{is} = \hat{V}_i + u_{is} \quad \forall i, s$$

To describe the rankings, we compute, for each simulation s , the average program composition at each rank decile, with higher deciles corresponding to better ranked students in a given program. Ultimately, we report the average of the decile averages across simulations. Panel (a) of Figures 2.7 to 2.9 displays the observed shares of female, low-SES, from high school, from district and top five cities students at each rank decile, while Panel (b) reports the differences between the observed and simulated baseline shares, for the three types of CPGE separately.

The reverse-engineered ranking rules from which we infer the CPGE preferences generate rankings of applicants that are overall similar to the observed rankings, but underestimate the degree of preference for applicants that are geographically-close at the top of the rankings: for all considered demographic characteristics and at all deciles except the last one, the observed and simulated decile compositions are not significantly different.¹⁵ This suggests that the model does a good job at predicting the observed rankings, except at the very top of the ranking, where CPGE programs seem to refine the raw score-based rankings of applicants, by giving an additional premium to geographically-close applicants (the high school and top 5 cities applicants for science CPGE, as well as district applicants for business and humanities CPGE).

Overall, our analysis suggests that programs rank applicants based on their achievement level in the core subjects of the considered specialization, but also based on their demographic characteristics, assigning moderately sized positive weights to low-income applicants, moderate negative (positive) weights to female applicants in female-dominated (male-dominated) specializations, and large positive weights to applicants from the considered high school or district.

2.4.3 Controlling for Predicted *Baccalauréat* Grades

Positive or negative weights on applicants' demographic characteristics may reflect an attempt from the part of the CPGE programs to contextualize applicants' continuous assessment grades, rather than a genuine preference for some demographic characteristics. Several reasons may make such

¹⁵The reported error bars measure the dispersion of the estimates across programs and the simulation error, but do not take into account the estimation error. As the estimation of a program's scoring model relies on only one ranking of applicants, i.e., one observation, we cannot use traditional asymptotics to compute the estimation error.

adjustments necessary. Raw continuous assessment grades are unlikely to be directly comparable across high schools, since both high school quality and grading schemes vary widely across high schools, or across students with different demographic characteristics, e.g. in the presence of grading biases favoring certain groups of students in certain subjects (Terrier, 2020; Lavy and Sand, 2018b; Carlana, 2019b).

To rule out the possibility that the positive/negative weights on applicants' demographic characteristics that we estimate using our baseline specification reflect such adjustments, we augment the model described by Equation (2.1) by adding applicants' predicted performance at the *baccalauréat* in each subject among the set of regressors. The prediction is performed separately for each combination of high school track, gender, and high vs. medium-high to low SES status, using applicants' continuous assessment grades, preliminary *baccalauréat* grades and a full set of fixed effects for high school of origin.¹⁶

Panel (b) of Table 2.5 reports the 25 percent lowest and strongest weights assigned to each of the applicants' demographic characteristics, when controlling for predicted *baccalauréat* grades.¹⁷ The introduction of predicted *baccalauréat* grades as controls has little effects on the magnitude of the weights assigned to the different demographic characteristics of applicants. For instance, for science CPGE, the premium attached to coming from the high school where the CPGE is located (respectively from the district) are reduced by 0.05 (0.04) standard deviation (from a baseline value of 0.30 (0.16) standard deviation), while weights on all other demographic characteristics are only marginally affected. Therefore, weights on applicants' demographic characteristics do not seem to be driven by CPGE program's attempt to contextualize their grades.

2.5 Counterfactual Simulations

The analysis of programs' preferences suggests that CPGE programs do not treat equally applicants with different demographic characteristics. To assess the amount of socio-demographic stratification generated by CPGE programs' preferences, we simulate the impact of a policy ban on applicants' demographic characteristics.

¹⁶The mean correlation between predicted and observed *baccalauréat* grades is of 0.87, with 95 percent of the subjects having a correlation value that lies between 0.82 and 0.92.

¹⁷Though the most natural statistic to report would be the median, we opt for the bottom and top quartile weights instead, as more than 50 percent of the CPGE programs do not take into account applicants' demographic characteristics when computing their scores.

2.5.1 Impact of a Ban on Applicants' Characteristics

To simulate how programs would adjust their ranking rules if they did not have access to information on applicants demographic characteristics, we reverse-engineer the observed programs' rankings of applicants based on the program-specific logistic lasso model described in Section 2.3, and using only the information on applicants' grades. To produce the counterfactual rankings, we let elite programs try and replicate their observed ranking of applicants using high school grades only, which assumes that programs have perfect information on the correlation between high school grades and the demographic characteristics that they value, and therefore provides the lower bound effect of the ban on elite programs rankings. We however consider this approach to be policy-relevant in the sense that in the limited information setting, programs would likely try and exploit the existing correlation between grades in a given subject and demographic characteristics to indirectly recover their preference-ranking of applicants.

Counterfactual rankings. Using the reverse-engineered weights on grades estimated in the limited-information setting, students are ranked based on their new grades-based score. Panel (c) of Figures 2.7 to 2.9 show the differences between the counterfactual and baseline shares of female, low-SES, from high school, from district and top five cities students at each decile of the rankings of applicants, separately for each CPGE type. When applicants are ranked based on grades only, the share of top five cities, high school and district applicants decreases by up to two percentage points in the upper half of the ranking, and increases by up to four percentage points among the 10 to 20 percent lowest ranked applicants, for all CPGE types.

Counterfactual assignments. To simulate the applicant-program match that the counterfactual rankings would generate, we run the APB matching mechanism using the simulated rankings of applicants for the CPGE programs while keeping constant the ranking of applicants for all other programs. Panel (a) of Table 2.8 reports the simulated baseline and counterfactual composition of the body of students admitted to CPGE programs. The table also shows, for comparison, the observed composition, which as expected is very close to the simulated baseline composition. For most applicants' characteristics except geographical origin, counterfactual program composition does not significantly depart from the simulated baseline composition. The policy ban, however, would significantly affect the geographical composition of the CPGE programs: the share of applicants

admitted to a CPGE located in their high school of origin decreases from 9.1 percent in the baseline simulations to 8.4 percent in the counterfactual simulations (a 7.7 percent decrease), while the share of top 5 cities students decreases from 9.4 percent to 9.1 percent (a 3.2 percent decrease).

Impact on offers received. Table 2.9 gives information on the impact of the ban on the offers received by the CPGE applicants ranked by at least one CPGE program. Compared to the baseline assignment, the counterfactual APB matching yields a different outcome for 4.8 percent of the CPGE applicants that were ranked by at least one CPGE program: 3.9 percent are assigned to a different program (hereafter designated as “movers”), 0.4 percent end up unassigned while they would have received an offer in the no-policy baseline and the same share of applicants receives an offer while they would have been unassigned at baseline. The ban on applicants’ demographic characteristics has only a limited impact on the rank of the program from which they receive an offer in the ROL they submitted to the platform, for most CPGE applicants profiles except the top five cities applicants. Among the movers, low-SES applicants tend to receive a better offer (-0.48 ranks in ROL), at the expense of their medium to high SES peers ($+0.08$ ranks in ROL). This effect is most likely explained by the impact of the ban on students from the top five cities, who incur a large negative effect ($+2.58$ ranks in ROL) to the benefit of students from other cities (-0.51 ranks).

2.5.2 Heterogeneous Effects of the Ban on Applicants’ Characteristics across Different Types of CPGE

CPGE programs are characterized by a high degree of qualitative differentiation, with a limited number of CPGE programs concentrating a large share of the students who are ultimately admitted to the top-tier *grandes écoles*, and the vast majority of other CPGE serving mostly second-tier *grandes écoles* or university graduate programs. First-tier CPGE programs are highly concentrated in urban areas, and in particular in the capital city, Paris. In this section, we investigate the impact of the policy ban on applicants’ characteristics on the composition of CPGE along these two dimensions of heterogeneity - selectivity and location.

By selectivity of CPGE programs. Table 2.D2 in Appendix 2.D investigates changes in the composition of CPGE programs by quartile of CPGE selectivity.¹⁸ Under the policy ban on applicants' characteristics, the composition of CPGE programs belonging to the different quartiles of selectivity is barely affected along the gender, SES and income dimensions. To the contrary, the counterfactual school and district of origin among admitted candidates significantly differs from the baseline composition for the 25 percent most selective CPGE: the share of top five cities students decreases from 12.1 to 11.1 percent, while the share of applicants from the same district as the CPGE decreases from 34.8 to 33.9 percent. At a more disaggregated level, the policy ban on applicants' demographic characteristics has very large effects on the share of students from the same high school as where the CPGE is located: in the most selective CPGE, their share decreases from 6.4 percent to 4.9 percent.

By location of CPGE programs. Table 2.D3 in Appendix 2.D shows that the counterfactual rankings of applicants generate important changes in the geographical origin of students admitted to a CPGE located in Paris, but do not affect any of the other dimensions of interest. The share of students from the top five cities in Parisian CPGE decreases from 29.7 percent to 28.2 percent, and the share of students from the high school where the CPGE is located decreases from 5.7 percent to 4.1 percent.

2.6 Discussion

In this section, we first assess the extent to which taking into account applicants' characteristics when making admissions decisions helps CPGE programs achieve their objective of maximizing the share of their students that access elite graduate schools (Section 2.6.1). Second, we discuss how our counterfactual simulations results might be affected by a reaction of students to the policy ban (Section 2.6.2).

¹⁸The quartiles are computed based on the pre-2016 CPGE students' mean percentile rank at the *baccalauréat* final grade, computed separately by high school track and year of examination.

2.6.1 Would the Policy Ban on Applicants' Characteristics Be Costly for CPGE Programs?

It may be objected that CPGE choices are rational, in the sense that the way in which they value applicants' characteristics could allow them to maximize the number of admitted students who ultimately access a top-tier elite graduate school. If this was the case, then the counterfactual policy that consists in putting a ban on applicants' demographic characteristics would be costly for CPGE programs, as it would prevent them from selecting the students with the highest expected probability of success.

To check whether this is the case, we compare the goodness of fit of logistic regressions that predict access to a top-tier elite graduate school either based on the applicants' baseline scores or based on their counterfactual scores. The regressions are performed separately for each CPGE, on the sample of ranked applicants. For each CPGE, we then compute the mean predicted probability of access to top-tier elite graduate schools, separately for applicants who ultimately accessed them, and those who did not.

Table 2.G1 in Appendix 2.G reports these two sets of probabilities averaged across programs.¹⁹ It can be noted that the baseline and counterfactual probabilities of accessing a top-tier *grande école* are highly similar, for both groups of students. For applicants who did not eventually access a top-tier *grande école*, their score-based probability of access is equal to 6.3 percent, both under the baseline and counterfactual approaches. For applicants who accessed a top-tier *grande école*, the mean access probability based on their counterfactual score (14.6 percent) is only marginally lower than that based on their baseline score (14.7 percent). Therefore, the ban on applicants' demographic characteristics does not appear to prevent CPGE programs from selecting the students with the highest expected probability of accessing the top-tier elite graduate schools, on average.

2.6.2 Effect of the Policy Ban when Students React to It

Students could react to the policy change by modifying their application strategies, or by investing differently in their academic preparation.

Although structural approaches often model students' application decisions based on their admission probabilities (Arcidiacono, 2005; Howell, 2010), the existing empirical evidence suggests

¹⁹The estimation is performed on the subsample of CPGE programs for which more than one percent of the applicants ultimately accessed a top-tier elite graduate schools.

only limited adjustments in the students' application patterns when they are confronted with a change in the admission criteria to post-secondary programs. Exploiting the 1996 to 1998 bans on race-based affirmative action in California and Texas, Card and Krueger (2005) and Antonovics and Backes (2013) find no evidence of adjustment among high achieving students belonging to racial minorities in their application to selective state institutions.

In the context of race-based affirmative action, some evidence has been brought on adjustments in students' pre-college investment. For instance, Akhtari et al. (2020) show that the reinstatement of affirmative action in Texas narrowed racial gaps in SAT scores, grades, and attendance. Such adjustments should however be limited in the French context. Unlike in countries explicitly implementing race-based affirmative action, little is known about the way French programs value students' demographic characteristics. It therefore appears unlikely that students anticipate differential admission chances based on their demographic characteristics, and invest in their college preparation accordingly.

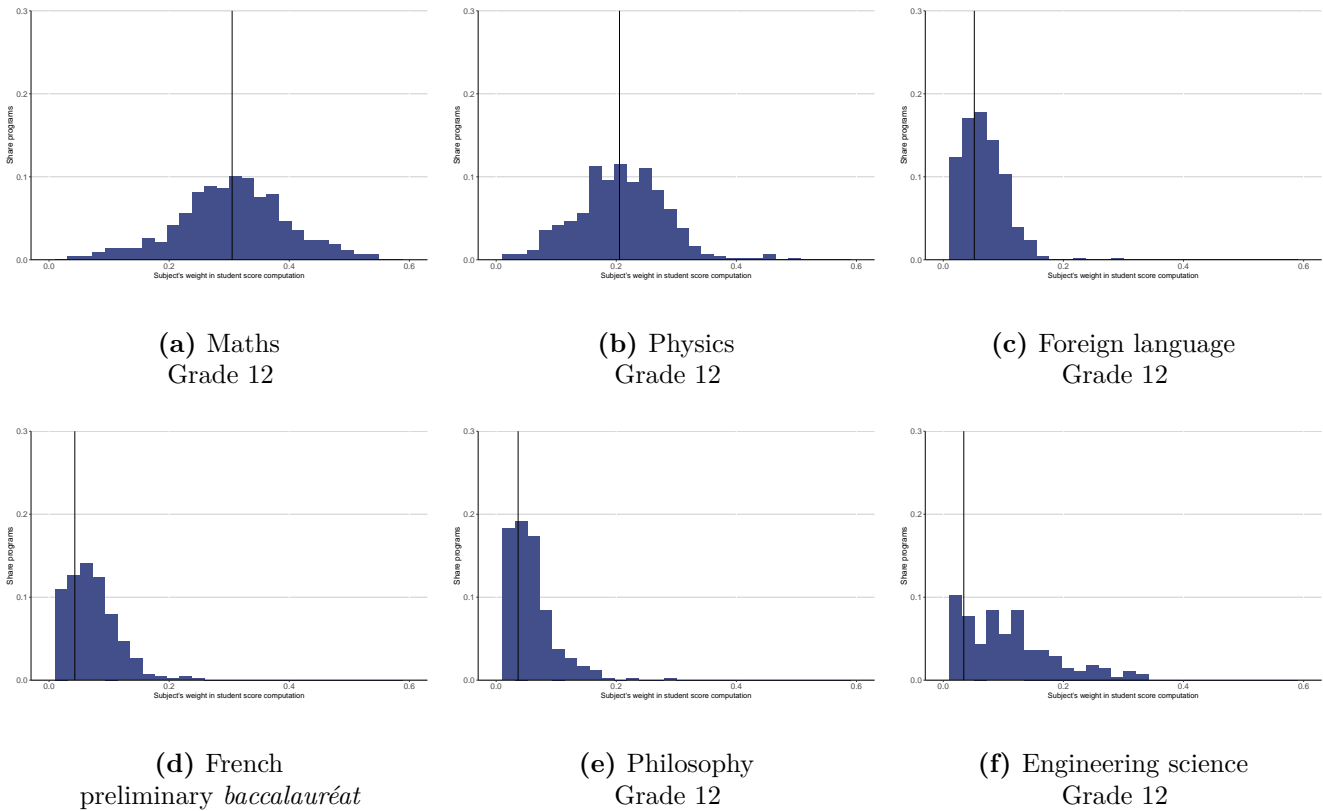
2.7 Conclusion

Leveraging rich data from the centralized system of college admissions in France, we estimate the preference of the most selective post-secondary programs – the CPGE – over applicants. Our results indicate that these programs give preferential treatment to geographically-close students and have a modest preference for gender balance: CPGE programs specializing in humanities tend to give preferential treatment to male students, whereas CPGE programs specializing in science tend to favor female students. We show that these differential treatments cannot be explained by an attempt from CPGE programs to contextualize applicants' grades.

We measure the contribution of the preferences of CPGE programs over applicants to the overall segregation by gender, socio-economic background and school and district of origin across programs by simulating the impact of a ban on applicants' characteristics when making admissions decisions. We find that the policy ban affects only moderately the rankings of applicants and the composition of admitted students to selective programs.

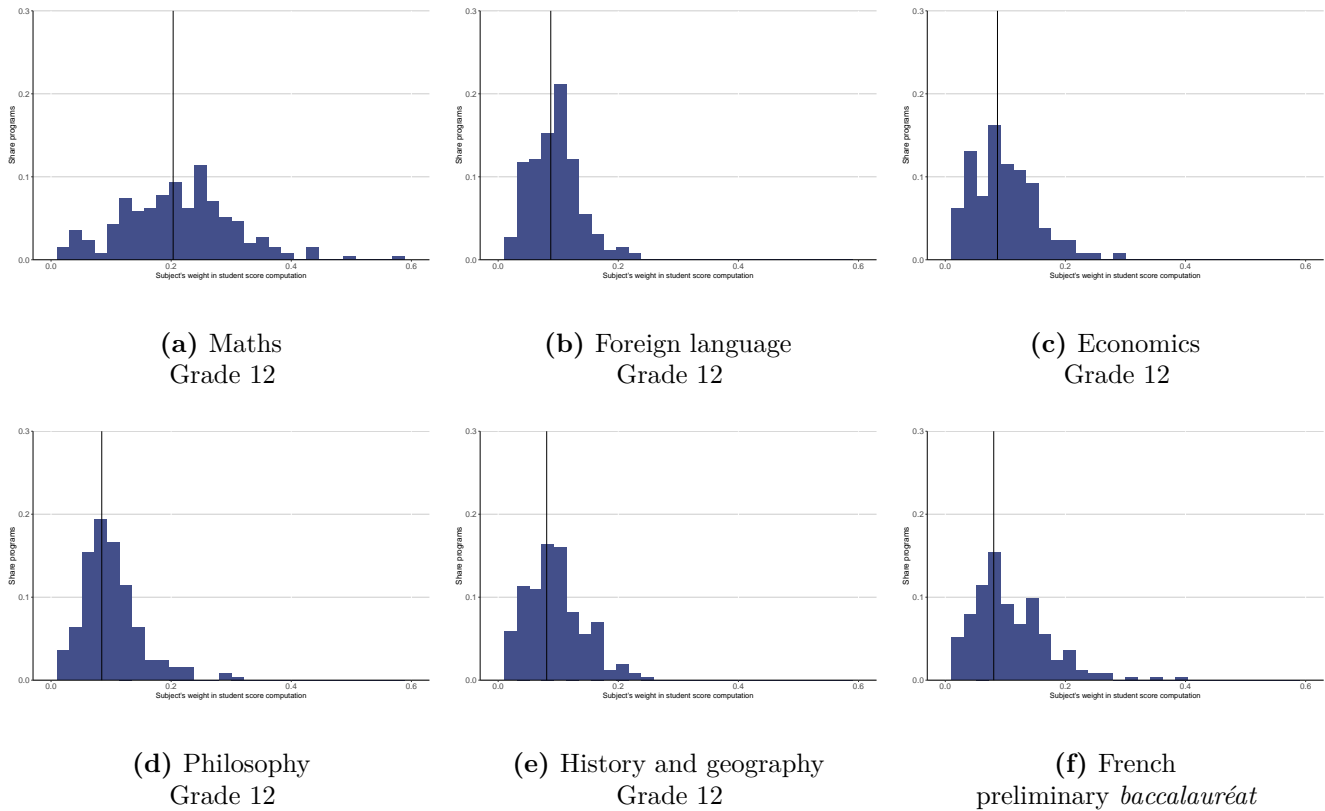
Overall, these findings suggest that the preferences of CPGE programs over applicants contribute little to the gender, social and spatial stratification of their student body. As a consequence, most of the observed stratification among CPGE programs must result from students' preferences—a research question that deserves to be investigated to complement the present analysis.

Figure 2.1 – Estimated Weights on Applicants’ Grades: Science CPGE



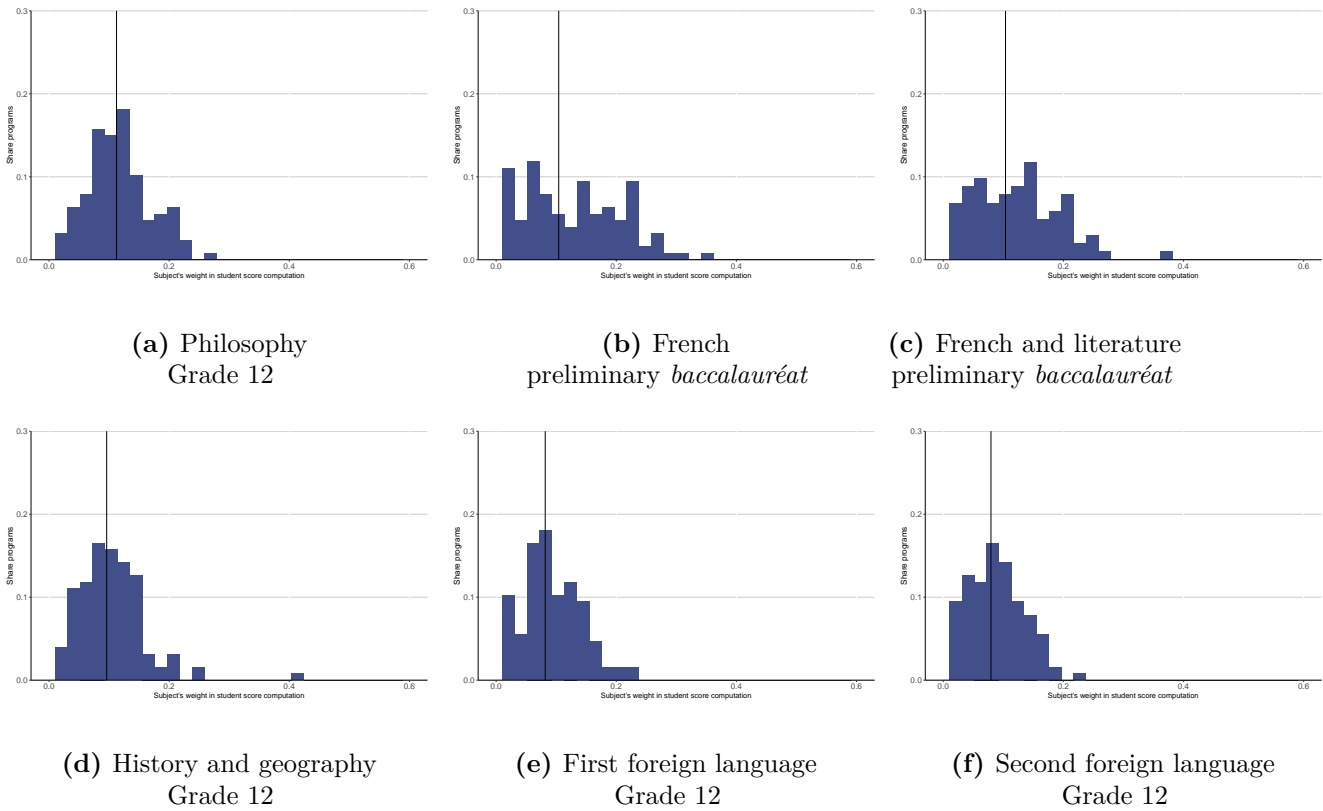
Notes: The figure plots the distribution across science CPGE programs of the coefficients on Grade 12 and anticipated *baccalauréat* grades obtained with the CPGE-specific logistic lasso estimation of the model described by Equation (2.1). The vertical line denotes the median coefficient across programs. The coefficients can be interpreted as the weights assigned to the considered subject in the applicants’ score computation underlying the programs’ rankings of applicants, where weights have been normalised to sum to one. The estimation is performed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Figure 2.2 – Estimated Weights on Applicants’ Grades: Business CPGE



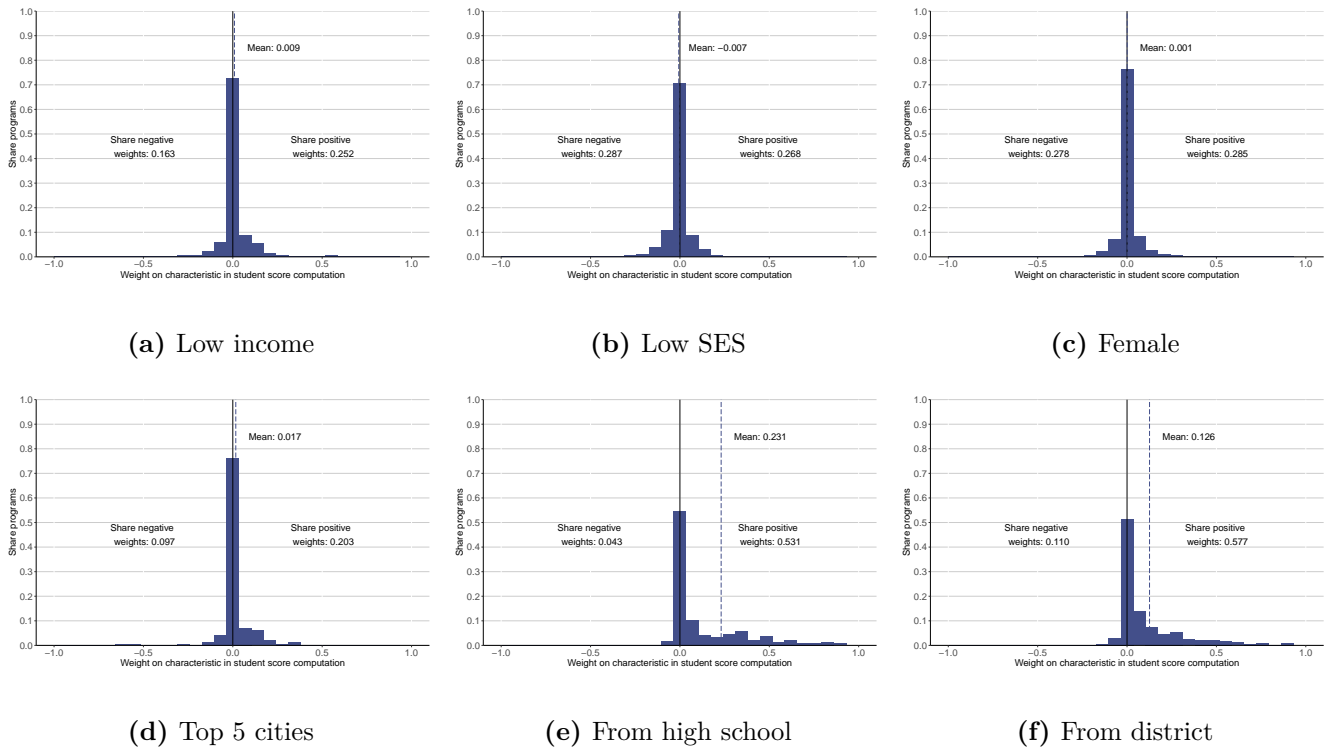
Notes: The figure plots the distribution across business CPGE programs of the coefficients on Grade 12 and anticipated *baccalauréat* grades obtained with the CPGE-specific logistic lasso estimation of the model described by Equation (2.1). The vertical line denotes the median coefficient across programs. The coefficients can be interpreted as the weights assigned to the considered subject in the score computation underlying the programs’ rankings of applicants, where weights have been normalised to sum to one. The estimation is performed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Figure 2.3 – Estimated Weights on Applicants’ Grades: Humanities CPGE



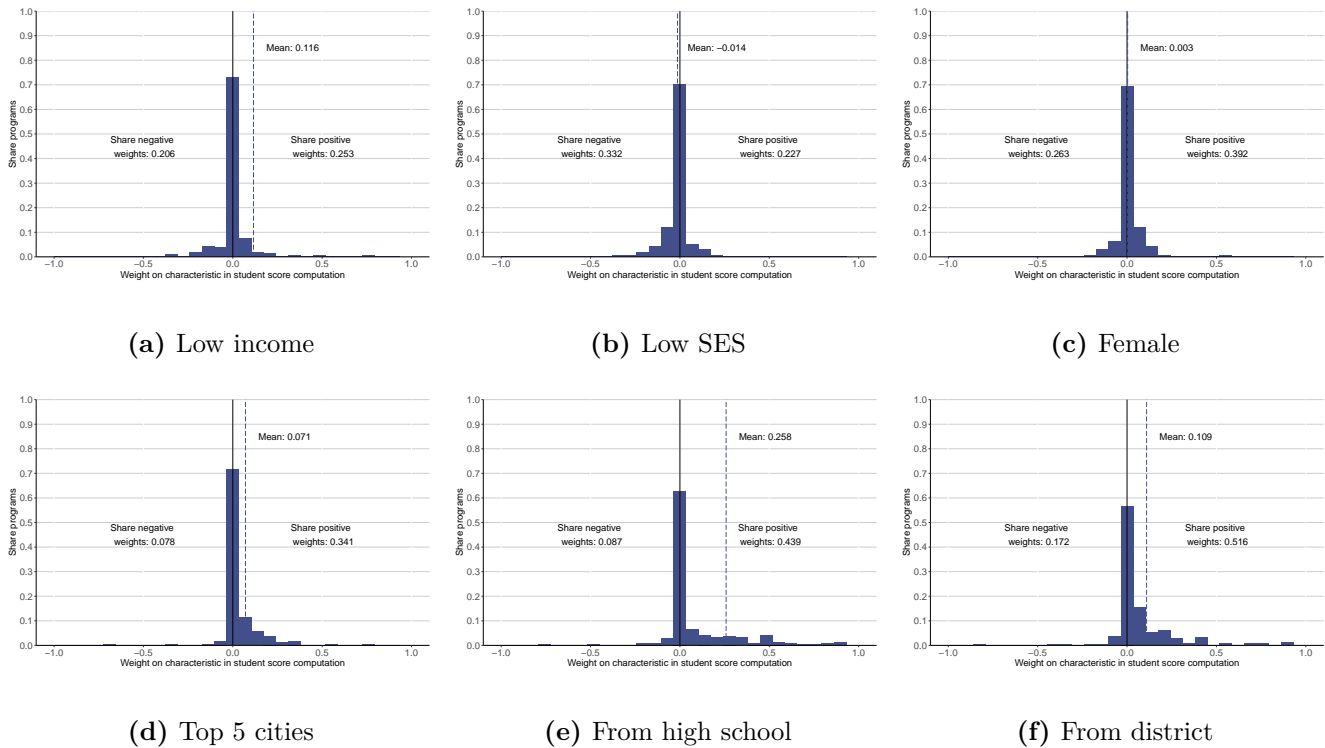
Notes: The figure plots the distribution across humanities CPGE programs of the coefficients on Grade 12 and anticipated *baccalauréat* grades obtained with the CPGE-specific logistic lasso estimation of the model described by Equation (2.1). The vertical line denotes the median coefficient across programs. The coefficients can be interpreted as the weights assigned to the considered subject in the score computation underlying the programs’ rankings of applicants, where weights have been normalised to sum to one. The estimation is performed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Figure 2.4 – Estimated Weights on Applicants’ Demographic Characteristics: Science CPGE



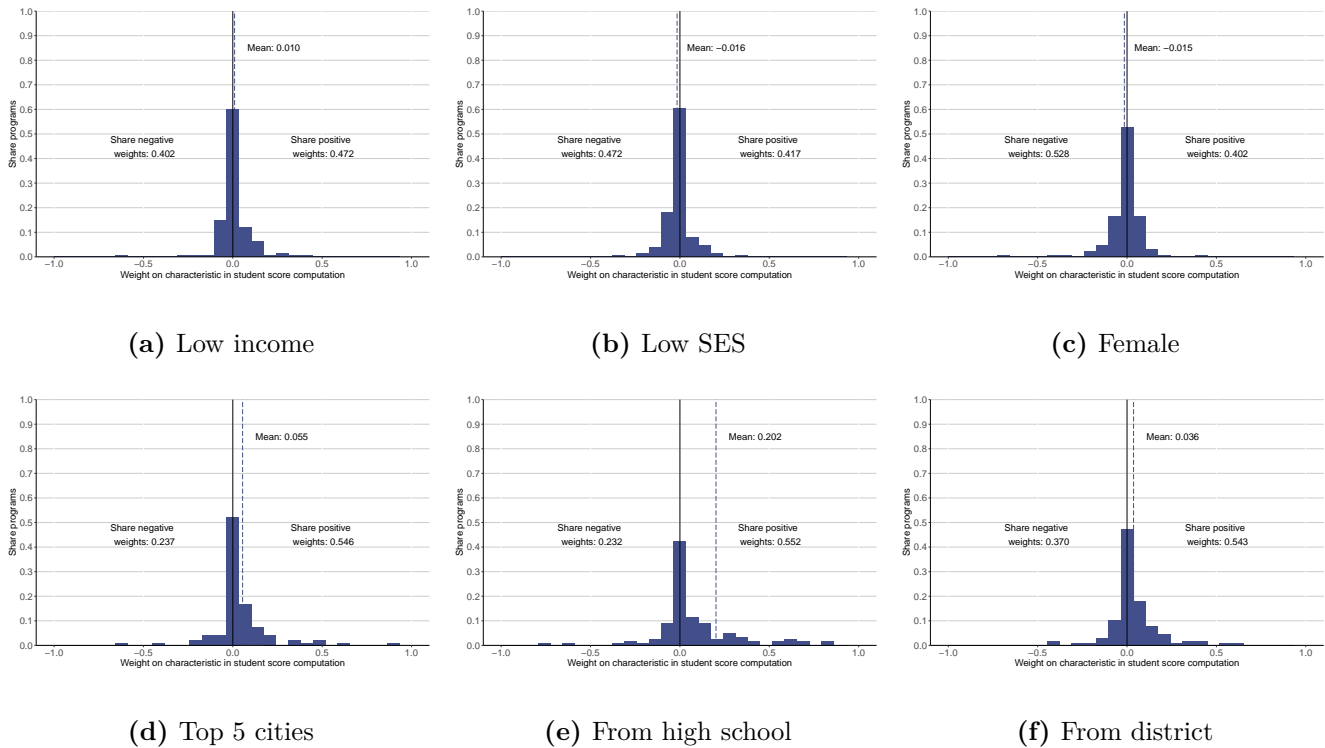
Notes: The figure plots the distribution across science CPGE programs of the coefficients on applicants’ demographic characteristics obtained with the CPGE-specific logistic lasso estimation of the model described by Equation (2.1). The vertical line denotes the median coefficient across programs. The coefficients can be interpreted as the bonuses/malus assigned to the considered characteristic in the score computation underlying the programs’ rankings of applicants. A bonus/malus of one can be interpreted as the equivalent of a one standard deviation increase/decrease in all the grades used in the computation of the grades-based score. The estimation is performed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Figure 2.5 – Estimated Weights on Applicants’ Demographic Characteristics: Business CPGE



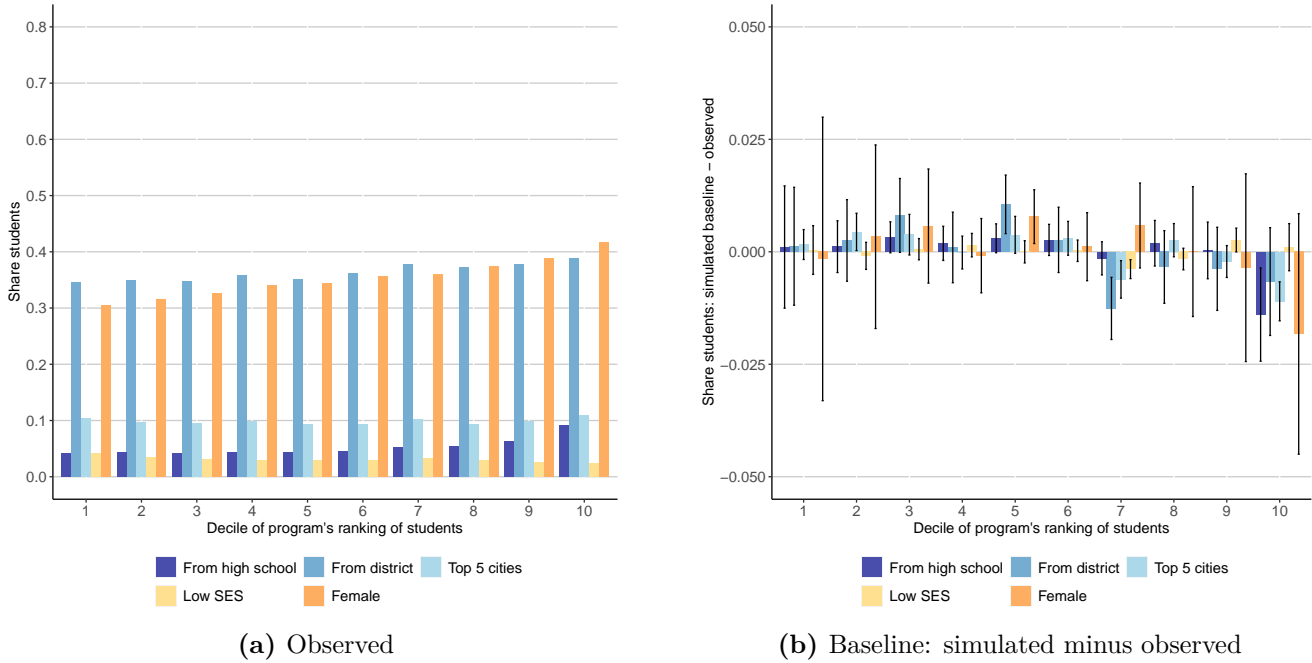
Notes: The figure plots the distribution across business CPGE programs of the coefficients on applicants’ demographic characteristics obtained with the CPGE-specific logistic lasso estimation of the model described by Equation (2.1). The vertical line denotes the median coefficient across programs. The coefficients can be interpreted as the bonuses/malus assigned to the considered characteristic in the score computation underlying the programs’ rankings of applicants. A bonus/malus of one can be interpreted as the equivalent of a one standard deviation increase/decrease in all the grades used in the computation of the grades-based score. The estimation is performed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Figure 2.6 – Estimated Weights on Applicants’ Demographic Characteristics: Humanities CPGE



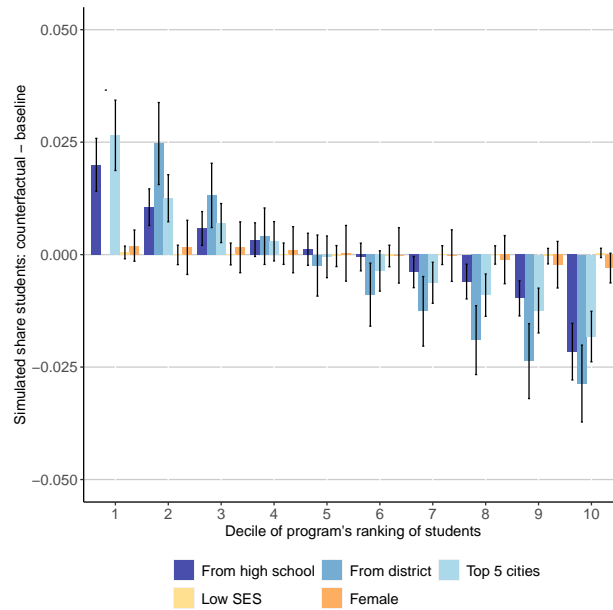
Notes: The figure plots the distribution across humanities CPGE programs of the coefficients on applicants’ demographic characteristics obtained with the CPGE-specific logistic lasso estimation of the model described by Equation (2.1). The vertical line denotes the median coefficient across programs. The coefficients can be interpreted as the bonuses/malus assigned to the considered characteristic in the score computation underlying the programs’ rankings of applicants. A bonus/malus of one can be interpreted as the equivalent of a one standard deviation increase/decrease in all the grades used in the computation of the grades-based score. The estimation is performed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Figure 2.7 – Applicants’ Characteristics by Decile of Programs’ Rankings: Science CPGE



(a) Observed

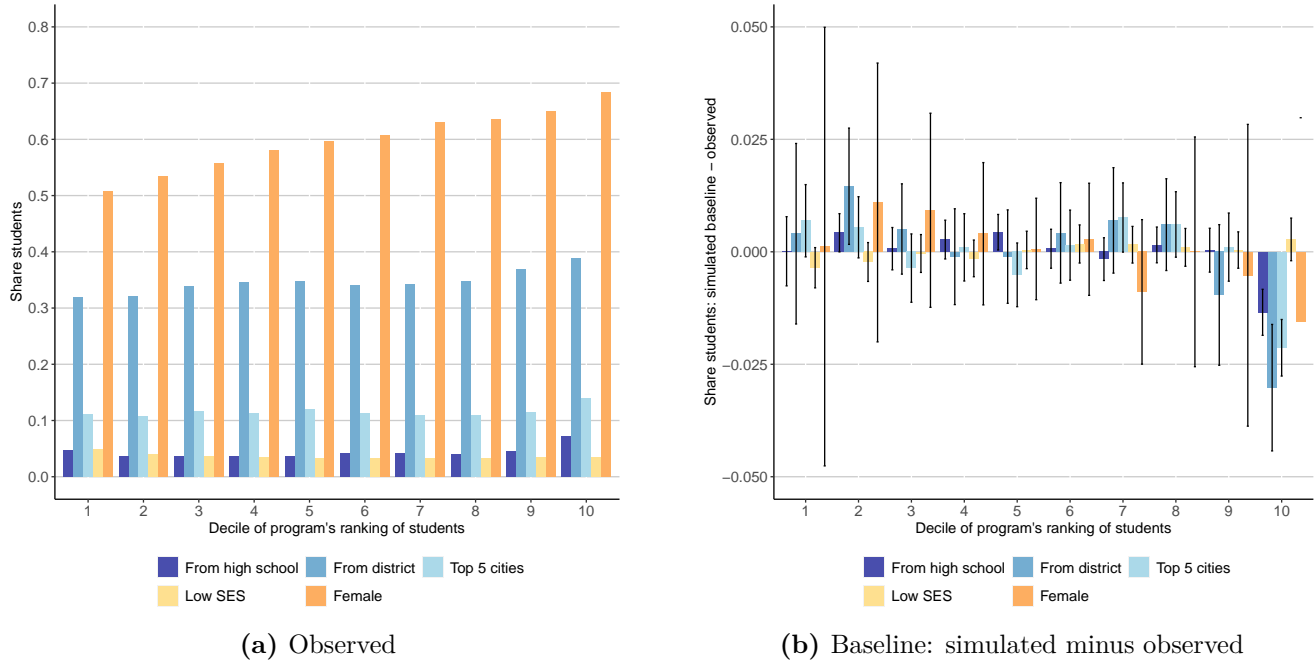
(b) Baseline: simulated minus observed



(c) Simulated counterfactual minus simulated baseline

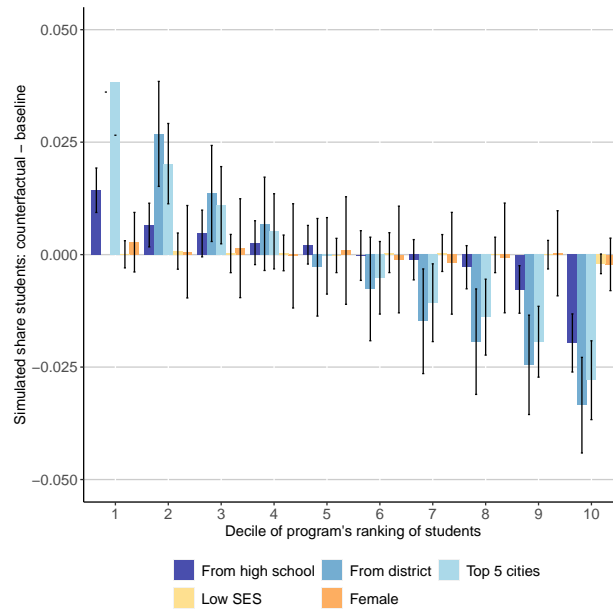
Notes: The figure plots the share of female, low-SES, top five cities, from district and from high school students at each decile of the science CPGE’s rankings of applicants (Panel (a)), the differences between baseline simulated and observed shares (Panel (b)) and the differences between simulated counterfactual and simulated baseline shares (Panel (c)). The simulated shares reported are averages over 300 simulations of the applicants’ scores and resulting rankings. In the baseline simulations, programs are allowed to compute applicants’ scores using the entire applicants’ profiles, while counterfactual simulations put a ban on applicants’ demographic characteristics. The estimation is performed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Figure 2.8 – Applicants’ Characteristics by Decile of Programs’ Rankings: Business CPGE



(a) Observed

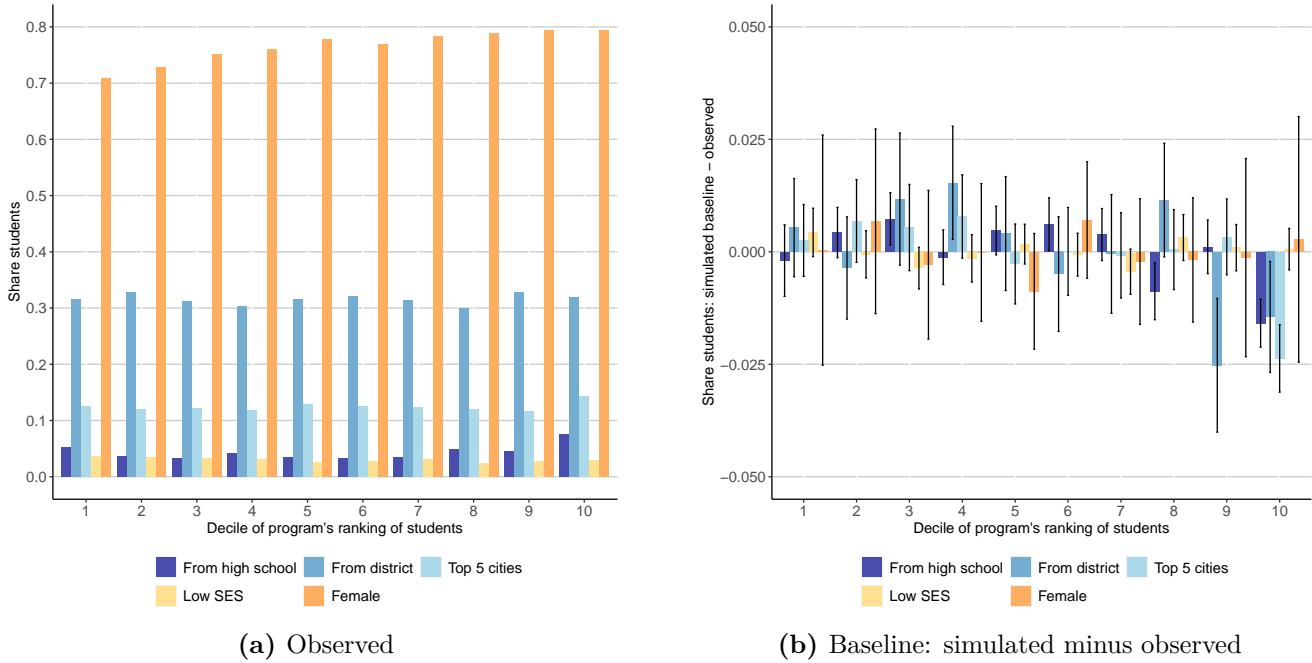
(b) Baseline: simulated minus observed



(c) Simulated counterfactual minus simulated baseline

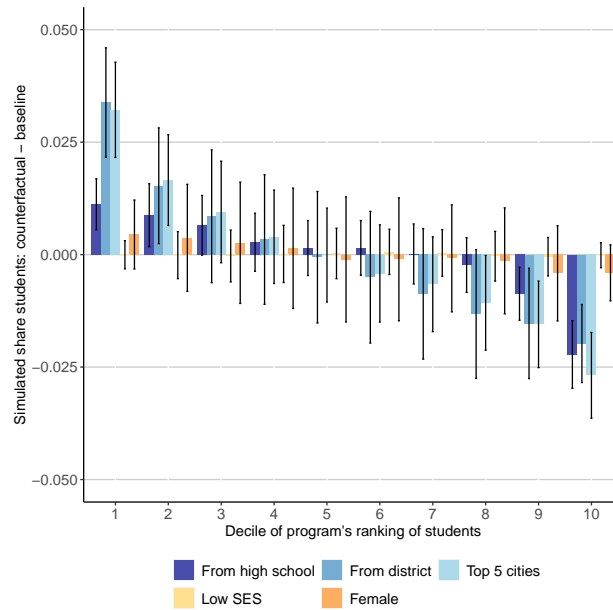
Notes: The figure plots the share of female, low-SES, top five cities, from district and from high school students at each decile of the business CPGE’s rankings of applicants (Panel (a)), the differences between baseline simulated and observed shares (Panel (b)) and the differences between simulated counterfactual and simulated baseline shares (Panel (c)). The simulated shares reported are averages over 300 simulations of the applicants’ scores and resulting rankings. In the baseline simulations, programs are allowed to compute applicants’ scores using the entire applicants’ profiles, while counterfactual simulations put a ban on applicants’ demographic characteristics. The estimation is performed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Figure 2.9 – Applicants’ Characteristics by Decile of Programs’ Rankings: Humanities CPGE



(a) Observed

(b) Baseline: simulated minus observed



(c) Simulated counterfactual minus simulated baseline

Notes: The figure plots the share of female, low-SES, top five cities, from district and from high school students at each decile of the humanities CPGE’s rankings of applicants (Panel (a)), the differences between baseline simulated and observed shares (Panel (b)) and the differences between simulated counterfactual and simulated baseline shares (Panel (c)). The simulated shares reported are averages over 300 simulations of the applicants’ scores and resulting rankings. In the baseline simulations, programs are allowed to compute applicants’ scores using the entire applicants’ profiles, while counterfactual simulations put a ban on applicants’ demographic characteristics. The estimation is performed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.1 – Summary Statistics: Applicants to CPGE Programs

	All (1)	Gender		SES		From top 5 cities	
		Male (2)	Female (3)	Low to Medium- High (4)	High (5)	No (6)	Yes (7)
Demographic Characteristics							
Female	0.47	-	-	0.48	0.47	0.47	0.48
Low income	0.11	0.10	0.11	0.25	0.03	0.11	0.10
From top 5 cities	0.09	0.09	0.10	0.06	0.11	-	-
High SES	0.64	0.64	0.63	-	-	0.63	0.76
Medium-high SES	0.13	0.13	0.13	0.36	0.00	0.14	0.07
Medium-low SES	0.17	0.17	0.17	0.47	0.00	0.18	0.12
Low SES	0.05	0.05	0.05	0.14	0.00	0.05	0.04
Academic Profile							
High school's average <i>baccalauréat</i> rank	52.08	51.86	52.31	48.15	54.29	51.50	57.58
High school track: science	0.66	0.74	0.57	0.59	0.70	0.66	0.64
High school track: business	0.18	0.13	0.24	0.18	0.19	0.18	0.21
High school track: humanities	0.07	0.03	0.13	0.08	0.07	0.07	0.09
High school track: other	0.08	0.10	0.07	0.15	0.05	0.09	0.06
Baccalauréat French grade	11.99	11.36	12.77	11.46	12.28	12.00	11.87
Baccalauréat philosophy grade	11.35	10.83	11.94	10.92	11.60	11.34	11.47
Baccalauréat foreign language grade	14.96	14.64	15.33	14.19	15.41	14.93	15.27
Baccalauréat maths grade	14.73	14.72	14.73	13.92	15.18	14.66	15.36
Baccalauréat physics grade	14.18	14.43	13.81	13.48	14.55	14.19	14.16
Applications							
Number of applications to CPGE	3.31	3.35	3.26	2.70	3.65	3.26	3.77
Applied to science CPGE	0.59	0.72	0.45	0.59	0.59	0.60	0.53
Applied to business CPGE	0.33	0.28	0.38	0.31	0.34	0.32	0.41
Applied to humanities CPGE	0.18	0.09	0.28	0.16	0.19	0.17	0.21
From same high school as CPGE	0.14	0.16	0.12	0.13	0.15	0.13	0.24
From same district as CPGE	0.77	0.77	0.78	0.73	0.80	0.79	0.57
Matching Outcomes							
Ranked by at least one CPGE	0.84	0.83	0.86	0.80	0.87	0.85	0.83
Ranked by all CPGE programs in ROL	0.40	0.37	0.43	0.44	0.37	0.41	0.27
Received an offer from CPGE	0.49	0.49	0.49	0.45	0.51	0.49	0.48
Rank of offer in ROL	3.23	3.49	2.95	2.79	3.45	3.24	3.21
N	87629	46096	41533	31712	55917	79380	8249

Notes: Each row reports the average of each variable listed in the first column, calculated on the universe of 2016 CPGE applicants as a whole (Column 1) and separately by gender (Columns 2–3), SES (Columns 4–5) and geographical origin (Columns 6–7). *Baccalauréat* grades are reported on a scale ranging from 0 to 20. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.2 – Summary Statistics: CPGE vs. Other Post-Secondary Programs

	Non-CPGE Programs	CPGE Programs			
		All	Science	Business	Humanities
	(1)	(2)	(3)	(4)	(5)
Programs' characteristics					
N programs	11435	836	438	267	131
Program capacity	219	58	62	48	65
% Programs offering joint admission to a dorm	-	77	85	63	81
% Programs using the decision-making tool to rank students	-	10	10	9	9
Applications received					
Number of applicants per program	422	501	594	389	420
Number of ranked applicants per program	298	297	354	213	276
Rank of last admitted student	271	217	250	160	226
Number of offers made per program	56	59	62	52	65
Ratio of number of applicants over program capacity	8	8	9	8	7
% Programs with vacant seats	45	29	28	28	36

Notes: The table gives basic information on the CPGE and other post-secondary programs' characteristics. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.3 – Summary Statistics: Ranked vs. Unranked Applications in CPGE

	Type of CPGE Program					
	Sciences		Business		Humanities	
	Unranked (1)	Ranked (2)	Unranked (3)	Ranked (4)	Unranked (5)	Ranked (6)
N applications	223485		94671		51067	
N applicants	51804		28691		15574	
% Applications	34.00	66.00	42.06	57.94	31.75	68.25
Demographic characteristics						
Female	0.34	0.32	0.50	0.59	0.68	0.76
Low income	0.10	0.10	0.09	0.11	0.09	0.10
From same high school as CPGE	0.03	0.09	0.02	0.07	0.02	0.07
From same district as CPGE	0.29	0.39	0.32	0.39	0.24	0.36
From top 5 cities	0.12	0.07	0.16	0.08	0.13	0.09
High SES	0.68	0.65	0.71	0.64	0.72	0.68
Medium-high SES	0.11	0.14	0.09	0.12	0.11	0.13
Medium-low SES	0.15	0.17	0.14	0.17	0.13	0.15
Low SES	0.05	0.04	0.05	0.06	0.03	0.04
Academic profile						
High school's average <i>baccalauréat</i> rank	53.61	51.27	54.89	52.88	53.42	52.38
High school track: science	0.96	0.86	0.46	0.37	0.32	0.27
High school track: economics	0.00	0.00	0.48	0.46	0.37	0.26
High school track: humanities	0.00	0.00	0.01	0.00	0.29	0.46
High school track: other	0.03	0.13	0.05	0.16	0.02	0.01
Baccalauréat French grade	11.15	11.97	11.43	12.64	13.58	14.27
Baccalauréat philosophy grade	10.58	11.15	10.99	11.81	12.44	13.29
Baccalauréat foreign language grade	14.47	14.94	14.71	15.24	15.63	16.29
Baccalauréat maths grade	14.07	15.26	14.43	15.31	14.47	14.33
Baccalauréat physics grade	13.77	14.93	12.57	13.52	12.69	13.15

Notes: Each row reports the average of the CPGE level averages of each variable listed in the first column for ranked vs. unranked applications, calculated separately by CPGE type. *Baccalauréat* grades are reported on a scale ranging from 0 to 20. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.4 – Summary Statistics: Applications by Quartile of Program’s Ranking

	Type of CPGE Program											
	Sciences				Business				Humanities			
	Student quartile in ranking				Student quartile in ranking				Student quartile in ranking			
	Q1 (1)	Q2 (2)	Q3 (3)	Q4 (4)	Q1 (5)	Q2 (6)	Q3 (7)	Q4 (8)	Q1 (9)	Q2 (10)	Q3 (11)	Q4 (12)
Demographic characteristics												
Female	0.29	0.31	0.33	0.37	0.51	0.57	0.61	0.66	0.71	0.77	0.78	0.80
Low income	0.10	0.10	0.10	0.09	0.12	0.11	0.10	0.11	0.11	0.10	0.10	0.09
From same high school as CPGE	0.07	0.08	0.08	0.11	0.07	0.06	0.07	0.08	0.07	0.07	0.06	0.08
From same district as CPGE	0.38	0.39	0.40	0.41	0.37	0.40	0.39	0.42	0.37	0.35	0.36	0.36
From top 5 cities	0.07	0.06	0.07	0.07	0.07	0.08	0.08	0.08	0.09	0.09	0.09	0.09
High SES	0.62	0.64	0.65	0.67	0.62	0.64	0.65	0.66	0.64	0.67	0.68	0.72
Medium-high SES	0.14	0.14	0.14	0.13	0.11	0.12	0.12	0.12	0.14	0.14	0.13	0.12
Medium-low SES	0.18	0.17	0.17	0.16	0.19	0.18	0.17	0.16	0.17	0.15	0.14	0.12
Low SES	0.05	0.04	0.04	0.04	0.06	0.06	0.05	0.06	0.05	0.04	0.04	0.03
Academic profile												
High school’s average <i>baccalauréat</i> rank	51.37	51.40	51.22	51.13	52.72	52.91	52.88	52.99	51.66	52.26	52.72	52.91
High school track: science	0.86	0.87	0.87	0.86	0.37	0.37	0.37	0.38	0.22	0.26	0.29	0.31
High school track: economics	0.00	0.00	0.00	0.00	0.46	0.47	0.46	0.45	0.29	0.28	0.27	0.22
High school track: humanities	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.47	0.45	0.44	0.47
High school track: other	0.13	0.13	0.13	0.13	0.16	0.16	0.16	0.17	0.01	0.01	0.00	0.00
Baccalauréat French grade	10.77	11.53	12.14	13.50	11.36	12.17	12.88	14.19	12.78	13.71	14.64	15.87
Baccalauréat philosophy grade	10.14	10.74	11.36	12.41	10.76	11.44	11.99	13.15	12.04	12.87	13.60	14.70
Baccalauréat foreign language grade	13.82	14.56	15.18	16.25	14.04	14.88	15.55	16.53	15.12	15.94	16.65	17.51
Baccalauréat maths grade	13.65	14.77	15.71	17.01	13.90	14.88	15.75	16.78	12.80	14.01	14.72	15.79
Baccalauréat physics grade	13.39	14.44	15.30	16.63	12.29	13.02	13.81	14.91	11.65	12.41	13.45	14.63

Notes: Each row reports the average of the CPGE level averages of each variable listed in the first column, computed separately by quartile of the CPGE rankings of applicants and CPGE specialisation. *Baccalauréat* grades are reported on a scale ranging from 0 to 20. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.5 – Reverse Engineered Weights on Applicants’ Characteristics

	Bottom Quartile Weight			Top Quartile Weight		
	Type of CPGE Program			Type of CPGE Program		
	Science (1)	Business (2)	Humanities (3)	Science (4)	Business (5)	Humanities (6)
Panel A. Main specification						
Female	-0.004	-0.002	-0.037	0.006	0.020	0.021
Low income	0.000	0.000	-0.015	0.000	0.000	0.029
Low SES	-0.007	-0.022	-0.036	0.003	0.000	0.012
From top 5 cities	0.000	0.000	0.000	0.000	0.053	0.065
From same high school	0.000	0.000	0.000	0.304	0.179	0.162
From same district	0.000	0.000	-0.008	0.163	0.102	0.068
From same <i>académie</i>	0.000	0.000	-0.006	0.088	0.050	0.043
Panel B. Controlling for predicted <i>baccalauréat</i> grades						
Female	0.000	-0.001	-0.043	0.017	0.015	0.019
Low income	0.000	0.000	-0.002	0.009	0.000	0.028
Low SES	-0.004	-0.013	-0.026	0.011	0.000	0.011
From top 5 cities	0.000	0.000	-0.000	0.000	0.022	0.054
From same high school	0.000	0.000	-0.000	0.250	0.118	0.140
From same district	0.000	0.000	-0.007	0.126	0.093	0.064
From same <i>académie</i>	0.000	0.000	-0.000	0.074	0.048	0.042
Panel C. Imputing missing grades						
Female	-0.007	-0.002	-0.033	0.006	0.016	0.020
Low income	0.000	0.000	-0.008	0.007	0.000	0.030
Low SES	-0.010	-0.017	-0.033	0.003	0.000	0.010
From top 5 cities	0.000	0.000	-0.000	0.000	0.045	0.072
From same high school	0.000	0.000	0.000	0.299	0.154	0.172
From same district	0.000	0.000	-0.004	0.158	0.103	0.077
From same <i>académie</i>	0.000	0.000	-0.011	0.094	0.050	0.054
N programs	429	258	127	429	258	127

Notes: For each demographic characteristic listed in the first column, each row reports the bottom and top quartile weights across CPGE, computed separately by CPGE type, and obtained from the CPGE-specific logistic lasso estimation of the model described by Equation (2.1). The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.6 – Characteristics of CPGE Programs Assigning Bottom Quartile Weight to Applicants' Characteristics

	Dep. var.: CPGE assigning bottom quartile weight to demographic characteristic:						
	Female	Low income	Low SES	From same high school	From same district	From same academic district	From top 5 cities
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Baseline CPGE	0.230 (0.036)	0.682 (0.037)	0.214 (0.035)	0.527 (0.041)	0.519 (0.038)	0.694 (0.040)	0.861 (0.048)
OLS coefficients:							
CPGE quartile of selectivity							
Second	-0.038 (0.047)	0.118** (0.049)	-0.090* (0.046)	-0.069 (0.054)	-0.012 (0.050)	-0.079 (0.053)	0.048 (0.061)
Third	0.015 (0.048)	0.022 (0.050)	0.017 (0.047)	-0.146*** (0.055)	-0.196*** (0.051)	-0.276*** (0.054)	-0.059 (0.060)
Fourth	0.058 (0.049)	0.085* (0.051)	0.065 (0.048)	-0.134** (0.056)	-0.290*** (0.052)	-0.283*** (0.055)	-0.086 (0.058)
CPGE located in Paris	0.055 (0.049)	-0.023 (0.051)	0.230*** (0.048)	-0.092 (0.058)	-0.251*** (0.052)	-0.209*** (0.054)	-0.281*** (0.050)
CPGE field of specialization							
Business	-0.005 (0.039)	-0.010 (0.040)	0.047 (0.038)	0.075* (0.045)	0.083** (0.041)	0.062 (0.043)	-0.121*** (0.045)
Humanities	0.152*** (0.047)	-0.213*** (0.049)	0.053 (0.046)	0.053 (0.054)	0.158*** (0.050)	0.028 (0.052)	-0.298*** (0.054)
N programs	698	697	698	674	698	672	498

Notes: Columns 1 to 7 report the OLS coefficients obtained from the regression on the CPGE characteristics listed in the first column of a dummy indicating whether the CPGE program assigns to the considered demographic characteristic a weight lower or equal to the bottom quartile weight. Standard errors are reported between parentheses. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.7 – Characteristics of CPGE Programs Assigning Top Quartile Weight to Applicants' Characteristics

	Dep. var.: CPGE assigning top quartile weight to demographic characteristic:						
	Female	Low income	Low SES	From same high school	From same district	From same academic district	From top 5 cities
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Baseline CPGE	0.153 (0.035)	0.272 (0.036)	0.290 (0.035)	0.149 (0.034)	0.195 (0.033)	0.151 (0.035)	0.097 (0.047)
OLS coefficients:							
CPGE quartile of selectivity							
Second	0.086* (0.046)	-0.087* (0.047)	0.067 (0.046)	0.043 (0.045)	-0.050 (0.043)	0.051 (0.046)	-0.013 (0.059)
Third	0.142*** (0.047)	0.025 (0.048)	-0.009 (0.047)	0.207*** (0.046)	0.167*** (0.044)	0.211*** (0.047)	0.085 (0.058)
Fourth	0.082* (0.048)	-0.035 (0.049)	-0.089* (0.048)	0.278*** (0.047)	0.210*** (0.045)	0.170*** (0.047)	0.117** (0.056)
CPGE located in Paris	0.039 (0.048)	0.014 (0.050)	-0.092* (0.048)	0.256*** (0.048)	0.337*** (0.045)	0.313*** (0.046)	0.235*** (0.048)
CPGE field of specialization							
Business	0.006 (0.038)	-0.015 (0.039)	-0.091** (0.038)	-0.068* (0.038)	-0.094*** (0.035)	-0.067* (0.037)	0.100** (0.044)
Humanities	0.048 (0.046)	0.132*** (0.047)	0.071 (0.046)	-0.170*** (0.045)	-0.204*** (0.043)	-0.168*** (0.045)	0.177*** (0.052)
N programs	698	697	698	674	698	672	498

Notes: Columns 1 to 7 report the OLS coefficients obtained from the regression on the CPGE characteristics listed in the first column of a dummy indicating whether the CPGE program assigns to the considered demographic characteristic a weight higher or equal to the top quartile weight. Standard errors are reported between parentheses. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.8 – Policy Ban on Applicants’ Characteristics: Impact on CPGE Composition

	Observed	Simulated	
	(1)	Baseline (2)	Counterfactual (3)
Panel A. Main specification			
Female	0.483	0.481	0.480
Low income	0.086	0.085	0.085
High SES	0.688	0.686	0.686
Medium-high SES	0.123	0.124	0.124
Medium-low SES	0.148	0.149	0.149
Low SES	0.036	0.036	0.037
From top 5 cities	0.094	0.094	0.091
From same high school as CPGE	0.097	0.091	0.084
From same district as CPGE	0.426	0.426	0.421
Panel B. Controlling for predicted <i>baccalauréat</i> grades			
Female	0.483	0.480	0.480
Low income	0.086	0.085	0.084
High SES	0.688	0.686	0.686
Medium-high SES	0.123	0.123	0.124
Medium-low SES	0.148	0.149	0.149
Low SES	0.036	0.036	0.036
From top 5 cities	0.094	0.094	0.092
From same high school as CPGE	0.097	0.091	0.085
From same district as CPGE	0.426	0.425	0.422
Panel C. Imputing missing grades			
Female	0.483	0.481	0.481
Low income	0.086	0.085	0.085
High SES	0.688	0.686	0.685
Medium-high SES	0.123	0.124	0.124
Medium-low SES	0.148	0.148	0.149
Low SES	0.036	0.036	0.037
From top 5 cities	0.094	0.094	0.091
From same high school as CPGE	0.097	0.092	0.085
From same district as CPGE	0.426	0.425	0.421
N	45380	45033	45001

Notes: For each variable listed in the first column, each row reports the observed, simulated baseline and simulated counterfactual average composition of admitted CPGE applicants. In the baseline simulations, programs are allowed to compute applicants’ scores using the entire applicants’ profiles, while counterfactual simulations put a ban on applicants’ socio-demographic profile. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.9 – Policy Ban on Applicants’ Characteristics: Impact on Matching Outcomes

	All	Female		Low SES		From top 5 cities	
	(1)	No (2)	Yes (3)	No (4)	Yes (5)	No (6)	Yes (7)
Aggregated effects							
Received different assignment	0.048 (0.001)	0.050 (0.002)	0.045 (0.001)	0.048 (0.001)	0.031 (0.002)	0.043 (0.001)	0.091 (0.002)
Assigned to a different program	0.039 (0.001)	0.040 (0.001)	0.038 (0.001)	0.040 (0.001)	0.026 (0.002)	0.035 (0.001)	0.078 (0.002)
Lost assignment to a program	0.004 (0.000)	0.004 (0.000)	0.004 (0.000)	0.004 (0.000)	0.003 (0.001)	0.004 (0.000)	0.009 (0.001)
Gained assignment to a program	0.004 (0.000)	0.005 (0.000)	0.004 (0.000)	0.004 (0.000)	0.003 (0.001)	0.004 (0.000)	0.004 (0.001)
ROL rank changes (movers only)							
Baseline average ROL rank	5.229 (0.050)	5.533 (0.084)	4.876 (0.050)	5.238 (0.051)	4.955 (0.139)	5.608 (0.067)	3.559 (0.105)
Change in ROL rank (counterfactual – baseline)	0.064 (0.046)	0.044 (0.089)	0.090 (0.083)	0.082 (0.049)	−0.475 (0.406)	−0.505 (0.059)	2.582 (0.150)
N	65397	34073	31324	62453	2944	59328	6069

Notes: For each variable listed in the first column, each row reports the average values and standard deviations (in parentheses) computed on the sample of CPGE applicants ranked at least once by a CPGE program. The reported averages are averages over 300 simulations, and the figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education. Standard errors are between parentheses.

Appendix to
How Do Elite Post-Secondary Programs
Select their Students?

2.A The APB Student-Program Matching

2.A.1 The traditional Gale Shapley Assignment Mechanism

The traditional Gale Shapley college-proposing deferred acceptance algorithm works as follows. In the first place, applicants submit their rank-order lists (ROL) of programs, while programs submit their ranking of applicants, which we designate hereafter as the “academic rankings”. In the first round of the procedure, each program j calls its n_j preferred applicants, where n_j corresponds to program j ’s capacity; all applicants are tentatively admitted to their preferred offer and reject all offers ranked below. In the following round, each program updates its ranking of applicants by removing the applicants who rejected its offer, and calls its new n_j preferred applicants. The procedure stops when no applicant rejects its offer.

2.A.2 The APB Matching Mechanism

The matching mechanism implemented on the APB platform deviates slightly from the Gale Shapley algorithm, as students can either apply to a program, or to a joint admission to a program and student accommodation provided by the program (generally in dorms), or to both options at the same time. While applicants’ priority in a given program is determined by the program-specific academic ranking of applicants, their priority at the dorm is determined by a dorm-specific social ranking of applicants, that depends on distance from the program and parental income.^{A.1} Importantly, the social rankings of applicants do not interfere with the academic rankings of applicants, since the academic and social rankings are managed by two different committees. The precise algorithm used by APB to achieve this dual matching has not been publicly disclosed, but the following assignment rule produces an applicant-program-dorm matching outcome that is very similar to the observed one (see Table 2.A1):

1. Each program j determines its pool of “eligible” applicants, which consists of (i) the applicants who applied to the program only, and (ii) the applicants who applied both to the program and the dorm that are ranked above the dorms’ capacity d_j in the social ranking of applicants.

^{A.1}The existence of two dimensions of student priority introduces some complexity in the assignment mechanism because the stability of the student assignment cannot be achieved on both dimensions at the same time: for any pair of students such that one student is ranked higher in the academic ranking but lower in the social ranking, choosing one over the other necessarily violates the stability of the assignment according to one of the two dimensions of student priority.

2. *Round 1.* Each program j calls its n_j preferred applicants within the pool of eligible applicants, where n_j is the program's capacity. The applicants are tentatively assigned to their preferred offer and reject all offers ranked below.
3. *Round k .* The programs update their rankings by removing applicants who received a better offer. The dorms update their social rankings by removing the applicants who received a better offer. Each program j calls its n_j preferred applicants within the new pool of eligible applicants. The applicants are tentatively admitted to their preferred offer and reject all offers ranked below.
4. The algorithm stops when no offer is rejected.
5. Some seats remain blocked by the program-and-dorms applicants who are ranked high on the social ranking but low on the academic ranking. These applicants cannot receive any offer because of their low academic rank, but prevent other program-and-dorms applicants that are better ranked on the academic ranking but lower ranked on the social ranking to join the pool of eligible applicants. These remaining seats are assigned by running the algorithm after manually removing the "blocking" applications.

2.A.3 Replication of the APB Assignment Mechanism

Table 2.A1 reports the share of applicants assigned to their observed assignment program under our interpretation of the APB assignment mechanism. Our procedure generates a matching outcome that is close to the observed one: 99.81 percent of applicants are assigned to their observed assignment program, with an accuracy rate that varies between 98.32 percent and 100.00 percent for the CPGE.

Table 2.A1 – Replication of the Gale Shapley Deferred Acceptance Algorithm

	Specialization	N pro-grams	% Correctly assigned students		
			All pro-grams	By program type	By program specialization
	(1)	(2)	(3)	(4)	(5)
CPGE programs					
Science	PC	139	99.81	98.69	98.32
	MP	126	99.81	98.69	98.88
	BCPST	56	99.81	98.69	99.00
	PT/PTSI/TB/TPC/TSI	119	99.81	98.69	98.81
	Other	6	99.81	98.69	99.12
Management	Economics major	101	99.81	99.38	99.38
	Technological major	43	99.81	99.38	99.85
	Sciences major	91	99.81	99.38	99.15
	Other	37	99.81	99.38	99.67
	Humanities	Classics	101	99.81	99.24
	Social sciences	26	99.81	99.24	99.18
	Other	7	99.81	99.24	100.00
Other programs					
	University program	3738	99.81	99.88	
	Undergraduate college	428	99.81	99.55	
	Undergraduate vocational and technical program	12032	99.81	99.88	
	Other	791	99.81	99.92	

Notes: The table presents the share of the 2016 APB applicants for whom our replication of the APB algorithm yields the same outcome as the observed one. The student assignment mechanism is simulated using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

2.B Applicants to CPGE - by CPGE Type

Table 2.B1 – Summary Statistics: Applicants to Science CPGE

	All	Gender		SES		From top 5 cities	
		Male	Female	Low to Medium- High	High	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Applicants' Demographic Characteristics							
Female	0.36	-	-	0.36	0.36	0.36	0.36
Low income	0.10	0.10	0.10	0.23	0.03	0.10	0.10
From top 5 cities	0.08	0.08	0.08	0.06	0.10	-	-
High SES	0.64	0.64	0.64	-	-	0.63	0.75
Medium-high SES	0.14	0.14	0.14	0.38	0.00	0.14	0.08
Medium-low SES	0.17	0.17	0.17	0.47	0.00	0.17	0.12
Low SES	0.05	0.05	0.05	0.14	0.00	0.05	0.04
Academic Profile							
High school's average <i>baccalauréat</i> rank	51.65	51.50	51.92	48.03	53.71	51.15	57.06
High school track: science	0.92	0.90	0.94	0.86	0.95	0.92	0.94
High school track: business	0.00	0.00	0.00	0.00	0.00	0.00	0.00
High school track: humanities	0.00	0.00	0.00	0.00	0.00	0.00	0.00
High school track: other	0.08	0.10	0.05	0.13	0.05	0.08	0.06
Baccalauréat French grade	11.76	11.24	12.70	11.27	12.04	11.78	11.56
Baccalauréat philosophy grade	10.98	10.61	11.64	10.52	11.25	10.96	11.20
Baccalauréat foreign language grade	14.80	14.57	15.20	14.03	15.24	14.78	15.03
Baccalauréat maths grade	14.74	14.84	14.56	13.92	15.21	14.68	15.35
Baccalauréat physics grade	14.45	14.65	14.10	13.67	14.90	14.45	14.55
Applications							
Number of applications to CPGE	3.44	3.49	3.35	2.82	3.79	3.38	4.07
From same high school as CPGE	0.15	0.17	0.12	0.15	0.16	0.14	0.25
From same district as CPGE	0.78	0.77	0.79	0.74	0.80	0.80	0.56
Matching Outcomes							
Ranked by at least one CPGE	0.85	0.85	0.84	0.80	0.88	0.85	0.85
Ranked by all CPGE programs in ROL	0.41	0.40	0.43	0.44	0.39	0.42	0.28
Received an offer from CPGE	0.49	0.51	0.45	0.44	0.51	0.49	0.50
Rank of offer in ROL	3.50	3.62	3.27	2.97	3.77	3.49	3.63
N	51804	33058	18746	18816	32988	47471	4333

Notes: Each row reports the average of each variable listed in the first column, calculated on the universe of 2016 science CPGE applicants as a whole (Column 1) and separately by gender (Columns 2–3), SES (Columns 4–5) and geographical origin (Columns 6–7). *Baccalauréat* grades are reported on a scale ranging from 0 to 20. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.B2 – Summary Statistics: Applicants to Business CPGE

	All	Gender		SES		From top 5 cities	
		Male	Female	Low to Medium- High	High	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Applicants' Demographic Characteristics							
Female	0.56	-	-	0.58	0.54	0.56	0.54
Low income	0.11	0.10	0.12	0.27	0.03	0.11	0.10
From top 5 cities	0.12	0.12	0.11	0.08	0.14	-	-
High SES	0.65	0.68	0.64	-	-	0.64	0.76
Medium-high SES	0.11	0.10	0.12	0.32	0.00	0.12	0.07
Medium-low SES	0.17	0.16	0.17	0.49	0.00	0.17	0.12
Low SES	0.06	0.05	0.06	0.16	0.00	0.06	0.04
Academic Profile							
High school's average <i>baccalauréat</i> rank	53.34	53.37	53.32	48.42	55.90	52.63	58.71
High school track: science	0.41	0.47	0.37	0.29	0.48	0.41	0.47
High school track: business	0.47	0.42	0.51	0.48	0.46	0.47	0.44
High school track: humanities	0.01	0.00	0.01	0.01	0.01	0.01	0.01
High school track: other	0.11	0.11	0.11	0.22	0.05	0.12	0.08
Baccalauréat French grade	12.14	11.58	12.59	11.56	12.44	12.15	12.08
Baccalauréat philosophy grade	11.48	11.09	11.78	11.06	11.70	11.47	11.52
Baccalauréat foreign language grade	14.96	14.70	15.18	14.08	15.45	14.91	15.38
Baccalauréat maths grade	14.81	14.55	15.03	13.98	15.26	14.72	15.48
Baccalauréat physics grade	13.22	13.26	13.18	12.39	13.49	13.20	13.34
Applications							
Number of applications to CPGE	2.54	2.41	2.63	2.12	2.75	2.53	2.61
From same high school as CPGE	0.12	0.13	0.11	0.10	0.12	0.11	0.19
From same district as CPGE	0.74	0.74	0.73	0.69	0.76	0.77	0.51
Matching Outcomes							
Ranked by at least one CPGE	0.79	0.75	0.83	0.76	0.82	0.80	0.74
Ranked by all CPGE programs in ROL	0.38	0.32	0.42	0.43	0.35	0.39	0.24
Received an offer from CPGE	0.47	0.45	0.49	0.42	0.50	0.48	0.42
Rank of offer in ROL	3.34	3.59	3.16	2.92	3.53	3.36	3.21
N	28691	12707	15984	9917	18774	25318	3373

Notes: Each row reports the average of each variable listed in the first column, calculated on the universe of 2016 business CPGE applicants as a whole (Column 1) and separately by gender (Columns 2–3), SES (Columns 4–5) and geographical origin (Columns 6–7). *Baccalauréat* grades are reported on a scale ranging from 0 to 20. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.B3 – Summary Statistics: Applicants to Humanities CPGE

	All	Gender		SES		From top 5 cities	
		Male	Female	Low to Medium- High	High	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Applicants' Demographic Characteristics							
Female	0.74	-	-	0.76	0.73	0.74	0.70
Low income	0.10	0.09	0.11	0.25	0.03	0.11	0.07
From top 5 cities	0.11	0.12	0.10	0.06	0.14	-	-
High SES	0.67	0.70	0.66	-	-	0.65	0.83
Medium-high SES	0.13	0.12	0.14	0.40	0.00	0.14	0.07
Medium-low SES	0.15	0.13	0.16	0.45	0.00	0.16	0.08
Low SES	0.04	0.04	0.04	0.13	0.00	0.04	0.02
Academic Profile							
High school's average <i>baccalauréat</i> rank	52.73	53.05	52.62	48.93	54.60	51.93	59.18
High school track: science	0.29	0.37	0.26	0.20	0.33	0.28	0.32
High school track: business	0.30	0.34	0.28	0.28	0.30	0.30	0.27
High school track: humanities	0.40	0.27	0.45	0.50	0.36	0.40	0.40
High school track: other	0.01	0.01	0.01	0.02	0.01	0.01	0.01
Baccalauréat French grade	14.00	13.58	14.19	13.64	14.13	14.04	13.65
Baccalauréat philosophy grade	12.95	12.88	12.97	12.60	13.11	12.98	12.70
Baccalauréat foreign language grade	16.08	15.90	16.15	15.48	16.38	16.05	16.31
Baccalauréat maths grade	14.82	14.45	15.00	14.08	15.11	14.72	15.63
Baccalauréat physics grade	13.45	13.57	13.39	12.72	13.67	13.38	13.98
Applications							
Number of applications to CPGE	2.77	2.61	2.82	2.42	2.94	2.72	3.12
From same high school as CPGE	0.12	0.12	0.12	0.12	0.12	0.11	0.22
From same district as CPGE	0.77	0.78	0.77	0.74	0.78	0.79	0.59
Matching Outcomes							
Ranked by at least one CPGE	0.84	0.79	0.86	0.82	0.85	0.84	0.82
Ranked by all CPGE programs in ROL	0.47	0.39	0.50	0.51	0.45	0.48	0.39
Received an offer from CPGE	0.60	0.58	0.60	0.57	0.61	0.60	0.60
Rank of offer in ROL	2.89	3.31	2.74	2.57	3.03	2.91	2.68
N	15574	4083	11491	5126	10448	13861	1713

Notes: Each row reports the average of each variable listed in the first column, calculated on the universe of 2016 humanities CPGE applicants as a whole (Column 1) and separately by gender (Columns 2–3), SES (Columns 4–5) and geographical origin (Columns 6–7). *Baccalauréat* grades are reported on a scale ranging from 0 to 20. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

2.C Computation Details on the Penalized Logit Model

This section details the derivation of the scoring model for each program in Section 2.3. Since the coefficients, $\boldsymbol{\gamma}$, on applicant characteristics are unconstrained, they do not add complexity to the analysis. Hereafter, we omit the contribution of characteristics to the scoring model, $\sum_{c=1}^C \gamma_{tc} d_{ic}$.

2.C.1 Model

For a given program, recall that applicant i 's score is determined as follows (after omitting $\sum_{c=1}^C \gamma_{tc} d_{ic}$):

$$S_i = \sum_{t=1}^T \mathbf{1}(T_i = t) \left[\sum_{k=1}^K \beta_{tk} \frac{\sum_{k'=1}^K \beta_{tk'}}{\sum_{k'=1}^K \beta_{tk'} \mathbf{1}(e_{ik'} = 1)} x_{ik} \right] + \epsilon_i \quad (\text{A.1})$$

such that

$$\begin{aligned} \beta_{tk} &\geq 0 \quad \forall t, k \\ \sum_{k=1}^K \beta_{tk} &> 0 \quad \forall t \end{aligned}$$

Without loss of generality, we index students by their position in the program's ranking, i.e., $i = 1$ is the student ranked first, $i = 2$ is the one ranked second, and so on.

Let $V_i = \sum_{t=1}^T \mathbf{1}(T_i = t) \left[\sum_{k=1}^K \beta_{tk} \frac{\sum_{k'=1}^K \beta_{tk'}}{\sum_{k'=1}^K \beta_{tk'} \mathbf{1}(e_{ik'} = 1)} x_{ik} \right]$. The log likelihood of observing the realized ranking of students, $i = 1, \dots, N$, is therefore:

$$\ln L(\boldsymbol{\beta}) = \ln \left[\prod_{i=1}^N \frac{\exp(V_i)}{\sum_{i' \geq i} \exp(V_{i'})} \right] = \sum_{i=1}^N \left(V_i - \ln \left[\sum_{i' \geq i} \exp(V_{i'}) \right] \right) \quad (\text{A.2})$$

We calculate the gradient and Hessian of the log likelihood function that will be useful in the estimation.

The gradient (score) with respect to β_{tp} , for all t and $p \in \{1, \dots, K\}$, can be written:

$$\begin{aligned} G_{tp}(\boldsymbol{\beta}) &= \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_{tp}} \\ &= \sum_{i=1}^N \mathbf{1}(T_i = t) \left(\frac{\beta^t}{\beta^{it}} x_{ip} + \psi_{itp} \sum_{k=1}^K \beta_{tk} x_{ik} + \frac{\sum_{i' \geq i} [\exp(V_{i'}) \mathbf{1}(T_{i'} = t) (\frac{\beta^t}{\beta^{i't}} x_{i'p} + \psi_{i'tp} \sum_{k=1}^K \beta_{tk} x_{i'k})]}{\sum_{i' \geq i} \exp(V_{i'})} \right) \\ &= \sum_{i=1}^N \mathbf{1}(T_i = t) \left(\xi_{itp} + \frac{\sum_{i' \geq i} [\exp(V_{i'}) \mathbf{1}(T_{i'} = t) \xi_{i'tp}]}{\sum_{i' \geq i} \exp(V_{i'})} \right) \end{aligned} \quad (\text{A.3})$$

where

$$\begin{aligned}\beta^t &= \sum_{l=1}^K \beta_{tl} \\ \beta^{it} &= \sum_{l=1}^K \beta_{tl} \mathbf{1}(e_{il} = 1) \\ \psi_{itp} &= \frac{1}{\beta^{it}} - \frac{\beta^t \mathbf{1}(e_{ip} = 1)}{(\beta^{it})^2} \\ \xi_{itp} &= \frac{\beta^t}{\beta^{it}} x_{ip} + \psi_{itp} \sum_{k=1}^K \beta_{tk} x_{ik}\end{aligned}$$

Similarly, the Hessian (with respect to β_{tp} and β_{sq} for all t and $p, q \in \{1, \dots, K\}$) can be written as:

$$\begin{aligned}H_{tp,sq}(\boldsymbol{\beta}) &= \frac{\partial^2 \ln L}{\partial \beta_{tp} \partial \beta_{sq}} \\ &= \sum_{i=1}^N \mathbf{1}(T_i = t) \left(+ \frac{\mathbf{1}(t = s) \left(\psi_{isq} x_{ip} + \psi_{itp} x_{iq} + \frac{\partial \psi_{itp}}{\partial \beta_{sq}} \sum_{k=1}^K \beta_{tk} x_{ik} \right) + \frac{\sum_{i' \geq i} \exp(V_{i'}) \mathbf{1}(T_{i'} = t) \mathbf{1}(t = s) \left[\xi_{i'tp} \xi_{i'sq} + \left(\psi_{i'sq} x_{i'p} + \psi_{i'tp} x_{i'q} + \frac{\partial \psi_{i'tp}}{\partial \beta_{sq}} \sum_{k=1}^K \beta_{tk} x_{i'k} \right) \right]}{\sum_{i' \geq i} \exp(V_{i'})} - \frac{\sum_{i' \geq i} \left[\exp(V_{i'}) \mathbf{1}(T_{i'} = t) \xi_{i'tp} \right] \times \sum_{i' \geq i} \left[\exp(V_{i'}) \mathbf{1}(T_{i'} = s) \xi_{i'sq} \right]}{\left(\sum_{i' \geq i} \exp(V_{i'}) \right)^2} \right) \end{aligned} \quad (\text{A.4})$$

where:

$$\frac{\partial \psi_{itp}}{\partial \beta_{sq}} = \begin{cases} \frac{2\beta^t \mathbf{1}(e_{ip}=1) \mathbf{1}(e_{iq}=1)}{(\beta^{it})^3} - \frac{\mathbf{1}(e_{ip}=1) + \mathbf{1}(e_{iq}=1)}{(\beta^{it})^2} & \text{if } s = t \\ 0 & \text{if } s \neq t \end{cases}$$

2.C.2 Penalized Conditional Logistic Regression (Lasso)

2.C.2.1 Penalized Model

We wish to find $\boldsymbol{\beta}$ that minimizes a penalized version of the (negative) log conditional likelihood (denoted $LL(\boldsymbol{\beta}; \lambda)$, using lasso (ℓ_1) penalty:

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} LL(\boldsymbol{\beta}; \lambda) = \arg \min_{\boldsymbol{\beta}} \left(-\ln L(\boldsymbol{\beta}) + \lambda \sum_{t=1}^T \sum_{k=1}^K |\beta_{tk}| \right)$$

where the parameter $\lambda \geq 0$ controls for the extent of regularization. As $\lambda \rightarrow \infty$, $\hat{\boldsymbol{\beta}}$ is eventually set to zero, while the maximum conditional likelihood estimate is obtained when $\lambda = 0$. As λ increases from 0, we obtain solutions with progressively fewer non-zero parameter estimates.

To enforce non-negativity of the coefficients $\boldsymbol{\beta}$, we adopt the following parametrization:

$$\beta_{tk} = (\theta_{tk})^2, \quad \forall t, k$$

With the reparametrized model, the above minimization problem becomes:

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta}} LL(\boldsymbol{\theta}^2; \lambda) = \arg \min_{\boldsymbol{\theta}} \left(-\ln L(\boldsymbol{\theta}^2) + \lambda \sum_{t=1}^T \sum_{k=1}^K (\theta_{tk})^2 \right) \quad (\text{A.5})$$

The penalized log-likelihood, the penalized gradient and the (non-penalized) hessian can be easily derived.

The penalized log-likelihood $LL(\boldsymbol{\theta}^2; \lambda)$ can be written as:

$$\begin{aligned} LL(\boldsymbol{\theta}^2; \lambda) &= -\ln L(\boldsymbol{\theta}^2) + \lambda \sum_{t=1}^T \sum_{k=1}^K (\theta_{tk})^2 \\ &= -\ln L(\boldsymbol{\beta}) + \lambda \sum_{t=1}^T \sum_{k=1}^K (\theta_{tk})^2 \end{aligned}$$

where $\ln L(\boldsymbol{\beta})$ is described in equation (A.2).

The gradient of the penalized log-likelihood $LL(\boldsymbol{\theta}^2; \lambda)$ is:

$$\begin{aligned} \frac{\partial LL(\boldsymbol{\theta}^2; \lambda)}{\partial \theta_{tp}} &= -\frac{\partial \ln L(\boldsymbol{\theta}^2)}{\partial \theta_{tp}} + 2\lambda \theta_{tp} \\ &= -\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_{tp}} \cdot \frac{\partial \beta_{tp}}{\partial \theta_{tp}} + 2\theta_{tp} \cdot \lambda \\ &= -2\theta_{tp} \cdot G_{tp}(\boldsymbol{\beta}) + 2\theta_{tp} \cdot \lambda, \quad \forall t, p \end{aligned}$$

where $G_{tp}(\boldsymbol{\beta})$ is given by equation (A.3).

2.C.2.2 Cyclic Coordinate Descent

To solve the constrained minimization problem (A.5), we employ the coordinate descent algorithm proposed by Friedman et al. (2010) and adapted to conditional logistic models by Reid and Tibshirani (2014).

The quadratic approximation of the log conditional likelihood $\ln L(\boldsymbol{\theta}^2)$ centered at point $\tilde{\boldsymbol{\theta}}$ is

$$\ln L(\boldsymbol{\theta}^2) = \ln L(\tilde{\boldsymbol{\theta}}^2) + (\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^\top G^{(\theta)}(\tilde{\boldsymbol{\theta}}^2) + \frac{1}{2}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^\top H^{(\theta)}(\tilde{\boldsymbol{\theta}}^2)(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})$$

where $G^{(\theta)}(\tilde{\boldsymbol{\theta}}^2) = \nabla_{\boldsymbol{\theta}} \ln L(\tilde{\boldsymbol{\theta}}^2)$ is the (TK) -vector of first derivatives (gradient) with respect to $\boldsymbol{\theta}$ and $H^{(\theta)}(\tilde{\boldsymbol{\theta}}^2) = \nabla_{\boldsymbol{\theta}}^2 \ln L(\tilde{\boldsymbol{\theta}}^2)$ is the Hessian.

The cyclic coordinate descent algorithm works as follows:

1. Initialize $\tilde{\boldsymbol{\theta}}$.
2. Compute $G^{(\theta)}(\tilde{\boldsymbol{\theta}}^2)$ and $H^{(\theta)}(\tilde{\boldsymbol{\theta}}^2)$.
3. Find $\hat{\boldsymbol{\theta}}$ minimizing $-(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^\top G^{(\theta)}(\tilde{\boldsymbol{\theta}}^2) - \frac{1}{2}(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}})^\top H^{(\theta)}(\tilde{\boldsymbol{\theta}}^2)(\boldsymbol{\theta} - \tilde{\boldsymbol{\theta}}) + \lambda \sum_{t=1}^T \sum_{k=1}^K (\theta_{tk})^2$.
4. set $\tilde{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}$.
5. Repeat Steps 2 to 4 until convergence.

The minimization in Step 3 above is performed one coordinate at a time. Letting $M(\boldsymbol{\theta}; \lambda)$ denote the objective in Step 3, we obtain the derivative:

$$\frac{\partial M(\boldsymbol{\theta}; \lambda)}{\partial \theta_{tp}} = -G_{tp}^{(\theta)}(\tilde{\boldsymbol{\theta}}^2) - (\theta_{tp} - \tilde{\theta}_{tp})H_{tp,tp}^{(\theta)}(\tilde{\boldsymbol{\theta}}^2) - \sum_{q \neq p} (\theta_{tq} - \tilde{\theta}_{tq})H_{tp,tq}^{(\theta)}(\tilde{\boldsymbol{\theta}}^2) + 2\lambda\theta_{tp}, \quad \forall t, p$$

Setting each of these equations to zero, we obtain the following update:

$$\hat{\theta}_{tp} = \frac{-\tilde{\theta}_{tp}H_{tp,tp}^{(\theta)}(\tilde{\boldsymbol{\theta}}^2) + G_{tp}^{(\theta)}(\tilde{\boldsymbol{\theta}}^2) + \sum_{q \neq p} (\hat{\theta}_{tq} - \tilde{\theta}_{tq})H_{tp,tq}^{(\theta)}(\tilde{\boldsymbol{\theta}}^2)}{-H_{tp,tp}^{(\theta)}(\tilde{\boldsymbol{\theta}}^2) + 2\lambda}, \quad \forall t, p$$

These are cycled over $p = 1, 2, \dots, K, 1, 2, \dots, K$ until convergence, each time updating the value of $\hat{\boldsymbol{\theta}}$ for use in subsequent updating equations. Note that $\tilde{\boldsymbol{\theta}}$, $G^{(\theta)}(\tilde{\boldsymbol{\theta}}^2)$ and $H^{(\theta)}(\tilde{\boldsymbol{\theta}}^2)$ remain fixed throughout this cyclic coordinate descent epoch.

Computation of $G^{(\theta)}$. The tp -coordinate of the gradient $G^{(\theta)}(\boldsymbol{\theta}^2)$ can be written:

$$\begin{aligned} G_{tp}^{(\theta)}(\boldsymbol{\theta}^2) &= \frac{\partial \ln L(\boldsymbol{\theta}^2)}{\partial \theta_{tp}} \\ &= \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_{tp}} \cdot \frac{\partial \beta_p}{\partial \theta_{tp}} \\ &= 2\theta_{tp} \cdot G_{tp}(\boldsymbol{\beta}), \quad \forall t, p \end{aligned}$$

where $G_{tp}(\boldsymbol{\beta})$ is given by equation (A.3).

Computation of $H^{(\theta)}$. The (tp, sq) -coordinate of the Hessian $H^{(\theta)}(\boldsymbol{\theta}^2)$ can be written:

$$\begin{aligned}
 H_{tp,sq}^{(\theta)}(\boldsymbol{\theta}^2) &= \frac{\partial^2 (\ln L(\boldsymbol{\theta}^2))}{\partial \theta_{tp} \partial \theta_{sq}} \\
 &= \frac{\partial}{\partial \theta_{sq}} \left(\frac{\partial \ln L(\boldsymbol{\theta}^2)}{\partial \theta_{tp}} \right) \\
 &= \frac{\partial}{\partial \theta_{sq}} \left(2\theta_{tp} \frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_{tp}} \right) \\
 &= 2\theta_{tp} \frac{\partial}{\partial \theta_{sq}} \left(\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_{tp}} \right) + 2 \left(\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \beta_{tp}} \right) \cdot \mathbf{1}_{\{tp=sq\}} \\
 &= 4\theta_{tp}\theta_{sq} \cdot H_{tp,sq}(\boldsymbol{\beta}) + 2G_{tp}(\boldsymbol{\beta}) \cdot \mathbf{1}_{\{sp=tq\}}
 \end{aligned}$$

where $H_{tp,sq}(\boldsymbol{\beta})$ is given by equation (A.4), and $G_{tp}(\boldsymbol{\beta})$ is given by equation (A.3).

2.D Results

2.D.1 Estimated Weights on Grades

Table 2.D1 – Reverse Engineered Weights on Grades

	Science		Business		Humanities	
	Share using subject (1)	Median weight (2)	Share using subject (3)	Median weight (4)	Share using subject (5)	Median weight (6)
Grade 11						
Maths	1.000	0.003	0.992	0.000	1.000	0.003
Physics	0.995	0.000	0.465	0.000	0.969	0.000
Earth and life science	0.876	0.000	0.465	0.000	0.953	0.000
Engineering science	0.641	0.000	0.016	0.000	0.000	-
Foreign language 1	1.000	0.000	0.992	0.000	1.000	0.010
Foreign language 2	0.995	0.000	0.992	0.000	1.000	0.011
Foreign language 3	0.235	0.000	0.589	0.000	0.134	0.000
History and geography	1.000	0.000	0.992	0.000	1.000	0.014
Latin	0.636	0.000	0.360	0.000	0.622	0.000
Social sciences	0.000	-	0.504	0.000	0.992	0.000
Literature	0.000	-	0.000	-	0.803	0.003
French	1.000	0.000	0.992	0.002	1.000	0.020
Sport	1.000	0.000	0.992	0.000	1.000	0.006
Sciences	0.012	0.000	0.504	0.000	0.992	0.004
Advanced foreign language	0.000	-	0.000	-	0.803	0.000
Maths - specialty	0.000	-	0.000	-	0.701	0.000
Foreign literature	0.012	0.000	0.000	-	0.803	0.000
Non-linguistic subject	0.867	0.000	0.822	0.000	0.835	0.000
Grade 12						
Maths	1.000	0.305	0.992	0.203	1.000	0.044
Physics	0.995	0.205	0.465	0.004	0.969	0.000
Earth and life science	0.876	0.003	0.465	0.000	0.953	0.000
Engineering science	0.639	0.034	0.012	0.000	0.000	-
Philosophy	0.995	0.037	0.981	0.084	1.000	0.112
Foreign language 1	1.000	0.052	0.992	0.088	1.000	0.082
Foreign language 2	0.995	0.000	0.992	0.067	1.000	0.079
Foreign language 3	0.219	0.000	0.558	0.000	0.134	0.000
History and geography	0.897	0.000	0.992	0.081	1.000	0.096
Social sciences	0.000	-	0.504	0.087	0.992	0.004
Literature	0.012	0.000	0.000	-	0.803	0.075
Sport	1.000	0.000	0.992	0.000	1.000	0.010
Advanced foreign language	0.000	-	0.000	-	0.803	0.007
Maths - specialty	0.865	0.019	0.829	0.002	0.843	0.014
Physics - specialty	0.804	0.007	0.314	0.005	0.181	0.016
Earth and life science - specialty	0.154	0.000	0.349	0.011	0.409	0.000
Foreign literature	0.012	0.044	0.000	-	0.803	0.005
Advanced economics	0.000	-	0.279	0.000	0.008	0.000
Social and political sciences	0.000	-	0.391	0.002	0.709	0.000
Computer science - specialty	0.207	0.010	0.019	0.000	0.000	-
Non-linguistic subject	0.865	0.000	0.829	0.000	0.850	0.000
Latin - culture	0.622	0.000	0.345	0.000	0.646	0.001
Anticipated <i>baccalauréat</i>						
French (oral)	0.995	0.023	0.981	0.056	1.000	0.054
French (written)	0.995	0.043	0.981	0.081	1.000	0.104
Interdisciplinary group project	0.876	0.000	0.829	0.000	1.000	0.008
Literature (oral)	0.012	0.000	0.000	-	0.803	0.057
Literature (written)	0.012	0.000	0.000	-	0.803	0.104
Biology	0.000	-	0.504	0.000	0.992	0.000
Science	0.012	0.000	0.000	-	0.803	0.000
Management	0.000	-	0.159	0.000	0.000	-

Notes: For each subject listed in the first column, each row reports the median weight obtained from the CPGE-specific logistic lasso estimation of the model described by Equation (2.1), and the share of programs for whom the weight is strictly positive. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

2.D.2 Heterogeneity by Program Type

Table 2.D2 – Counterfactual Assignment Simulations: CPGE Composition by Quartile of Selectivity

	Observed				Simulated							
					Baseline				Counterfactual			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Female	0.330	0.433	0.525	0.544	0.328	0.428	0.519	0.545	0.328	0.429	0.518	0.543
Low income	0.126	0.091	0.080	0.069	0.129	0.092	0.078	0.068	0.131	0.092	0.077	0.068
High SES	0.586	0.633	0.694	0.759	0.585	0.632	0.692	0.755	0.585	0.633	0.692	0.752
Medium-high SES	0.145	0.142	0.125	0.101	0.141	0.144	0.127	0.102	0.141	0.144	0.126	0.103
Medium-low SES	0.194	0.177	0.148	0.112	0.197	0.177	0.148	0.113	0.197	0.176	0.149	0.114
Low SES	0.067	0.042	0.029	0.025	0.070	0.041	0.029	0.026	0.070	0.041	0.029	0.026
From top 5 cities	0.088	0.058	0.088	0.123	0.093	0.058	0.087	0.121	0.096	0.059	0.085	0.111
From same high school as CPGE	0.140	0.124	0.081	0.075	0.143	0.122	0.075	0.064	0.143	0.120	0.071	0.049
From same district as CPGE	0.518	0.469	0.440	0.351	0.520	0.471	0.440	0.348	0.521	0.474	0.434	0.339
N	6973	8566	11497	14297	8087	11735	17909	19765	8097	11832	18151	19971

Notes: For each variable listed in the first column, each row reports the observed, simulated baseline and simulated counterfactual average composition of admitted CPGE applicants, separately by CPGE quartile of selectivity. In the baseline simulations, programs are allowed to compute applicants' scores using the entire applicants' profiles, while counterfactual simulations put a ban on applicants' socio-demographic profile. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

Table 2.D3 – Counterfactual Assignment Simulations: CPGE Composition by Geographical Location

	Observed		Simulated			
	RoF (1)	Paris (2)	Baseline		Counterfactual	
			RoF (3)	Paris (4)	RoF (5)	Paris (6)
Female	0.476	0.510	0.474	0.509	0.474	0.508
Low income	0.089	0.072	0.089	0.071	0.088	0.069
High SES	0.664	0.792	0.662	0.791	0.661	0.790
Medium-high SES	0.135	0.070	0.136	0.071	0.136	0.071
Medium-low SES	0.158	0.106	0.158	0.106	0.159	0.107
Low SES	0.039	0.025	0.039	0.025	0.039	0.025
From top 5 cities	0.047	0.301	0.048	0.297	0.047	0.282
From same high school as CPGE	0.104	0.066	0.099	0.057	0.094	0.041
From same district as CPGE	0.457	0.290	0.457	0.290	0.452	0.285
N	33723	7610	39169	11012	39354	11167

Notes: For each variable listed in the first column, each row reports the observed, simulated baseline and simulated counterfactual average composition of admitted CPGE applicants, separately for Parisian vs. Rest of France (RoF) CPGE. In the baseline simulations, programs are allowed to compute applicants' scores using the entire applicants' profiles, while counterfactual simulations put a ban on applicants' socio-demographic profile. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

2.E Goodness of Fit for Programs Using the DMT to Rank Applicants

Table 2.E1 reports the correlation coefficients between the observed and simulated baseline rankings of applicants, computed separately for the CPGE programs using the DMT to rank applicants and those not using it. Since the presence of error in the model is necessary to the estimation of the model described by Equation (2.1), the comparison can only be performed between the programs that do not use the DMT to rank applicants and the programs that use the DMT but refine ex-post the raw rankings produced by the DMT. As expected, we find that the correlation between the observed and simulated baseline rankings of applicants is higher for the programs using the DMT (with correlation coefficients ranging from 0.89 in business CPGE to 0.95 in science CPGE) compared to those not using it to rank applicants (with correlation coefficients ranging from 0.88 in humanities CPGE to 0.91 in science CPGE). However, both sets of coefficients do not differ significantly, suggesting that the scoring approach described in Section 2.3 is a good approximation to the behavior of the CPGE programs who do not rely on the DMT to rank applicants.

Table 2.E1 – Correlation Between Observed and Simulated Baseline Rankings of Applicants

	Uses DMT to rank applicants	
	No (1)	Yes (2)
Science CPGE	0.913	0.950
Business CPGE	0.894	0.886
Humanities CPGE	0.875	0.943
N CPGE programs	737	77

Notes: For each CPGE type listed in the first column, each row reports the correlation between observed and simulated baseline rankings of applicants, computed separately for CPGE programs using the DMT to rank applicants and those not using it. When CPGE programs use the DMT without refining the rankings of applicants in a second step, the logit estimation procedure fails, due to the absence of error in the model. Therefore, among DMT users, only those who are not using the raw DMT rankings are reported in the table. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

2.F Dealing with Missing Values on Grades

Table 2.F1 presents the share of students with non-missing grades for each subject in Grades 11 and 12, as well as for the *baccalauréat* preliminary exams, computed separately for the five high school tracks. For the core subjects of the considered track, the share of students with missing grades rarely exceeds one percent. Missing values on grades can however be numerous when, for a given high school track, several elective courses can be undertaken by the students. In this case, the share of non-missing values can be as low as, for instance, 15 percent for science track students in Latin.

As missing values mainly reflect the elective course choices of the students, our main approach assumes that CPGE programs systematically overlook the subjects with missing grades when computing the applicants' scores. Alternatively, it could be the case that the CPGE try to impute missing grades, when the considered subjects are part of the core subjects of the considered high school track.

Panel (c) of Table 2.5 reports the estimated weights on applicants' characteristics when missing values on the core subjects are imputed.^{A.1} For all CPGE types, the weights obtained are very similar to the weights obtained with the main specification. Therefore, the resulting baseline and counterfactual programs' sociodemographic composition are highly similar to those obtained using the main specification (see Panel (c) of Table 2.8).

^{A.1}For each high school track, we consider as core subjects those for which the grades are missing for less than 10 percent of the applicants. Missing values in these subjects are imputed using a mean-matching procedure that relies on the grades obtained in all other core subjects. For simplicity, missing values have been imputed based on the entire data, rather than separately for each CPGE.

Table 2.F1 – Share Students with Non-Missing Values on Grades

	High school curriculum				
	Sciences (1)	Economics (2)	Humanities (3)	Technical (4)	Other (5)
Grade 11					
Maths	0.995	0.994	0.039	0.994	0.882
Physics	0.995	0.023	0.026	0.555	0.496
Earth and Life science	0.873	0.022	0.025	0.006	0.367
Foreign Language 1	0.989	0.987	0.985	0.994	0.870
Foreign Language 2	0.993	0.992	0.990	0.925	0.768
History and geography	0.994	0.994	0.992	0.993	0.858
Latin	0.154	0.110	0.195	0.000	0.069
Social sciences	0.000	0.972	0.012	0.001	0.079
French	0.995	0.994	0.981	0.993	0.774
Sport	0.982	0.975	0.967	0.971	0.960
Sciences	0.000	0.969	0.964	0.002	0.122
Foreign Literature	0.000	0.000	0.938	0.000	0.089
Non-linguistic subject	0.259	0.234	0.233	0.044	0.133
Grade 12					
Maths	0.997	0.996	0.000	0.998	0.840
Physics	0.997	0.000	0.000	0.555	0.461
Earth and Life science	0.876	0.000	0.000	0.000	0.338
Philosophy	0.998	0.999	0.997	0.998	0.702
Foreign Language 1	0.994	0.991	0.990	0.998	0.871
Foreign Language 2	0.996	0.996	0.994	0.915	0.693
History and geography	0.997	0.998	0.997	0.463	0.763
Social sciences	0.000	0.999	0.000	0.000	0.062
Literature	0.000	0.000	0.996	0.000	0.058
Sport	0.901	0.904	0.858	0.911	0.861
Maths - specialty	0.545	0.613	0.244	0.000	0.225
Foreign Literature	0.000	0.000	0.992	0.000	0.058
Non-linguistic subject	0.264	0.238	0.234	0.047	0.095
Anticipated <i>baccalauréat</i>					
French (oral)	0.998	0.999	0.000	0.979	0.577
French (written)	0.998	0.999	0.000	0.979	0.577
Interdisciplinary Group project	0.998	0.999	0.998	0.000	0.651
Literature (oral)	0.000	0.000	0.998	0.000	0.093
Literature (written)	0.000	0.000	0.998	0.000	0.093
Biology	0.000	0.999	0.000	0.000	0.079
History and Geography	0.000	0.000	0.000	0.535	0.015
Science	0.000	0.000	0.998	0.000	0.093

Notes: For each subject listed in the first column, each row reports the share of CPGE applicants with non-missing grades, computed separately by high school curriculum. The figures are computed using data from the 2016 centralized undergraduate admissions procedure in France (APB), managed by the French Ministry of Higher Education.

2.G Would the Policy Ban on Applicants' Characteristics Be Too Costly for CPGE Programs?

Table 2.G1 – Predicted Probability of Access to a Top-Tier Grande École

	Predicted probability		N programs (3)
	Baseline score (1)	Counterfactual score (2)	
Panel A. Applicants who did not access top-tier grande école			
Mean probability of access	0.063	0.063	587
Panel B. Applicants who accessed top-tier grande école			
Mean probability of access	0.147	0.146	587

Notes: The table presents the mean predicted probabilities of access to top-tier *grandes écoles* within four years after applying to a CPGE, computed separately for applicants who accessed top-tier *grandes écoles* and applicants who did not. The probabilities are estimated using a CPGE-specific logistic regression based on the baseline (Column 1) or counterfactual (Column 2) scores computed by the CPGE to rank applicants, for the CPGE for which more than one percent of applicants accessed a top-tier *grande école*. In the baseline score computation, programs are allowed to use the entire applicants' profiles, while counterfactual computations are based on grades only. The estimation is performed using data from the 2016 centralized undergraduate admissions procedure in France (APB) and administrative enrollment data, managed by the French Ministry of Higher Education.

Chapter 3

Gendered Teacher Feedback, Students' Math Performance and Enrollment Outcomes: A Text Mining Approach *

We propose a novel approach to measure the degree of gender differentiation in the vocabulary that teachers use to assess their students' work. We build a model that predicts each student's gender based on the words appearing in the feedback she receives. For every teacher, we estimate a synthetic measure of gender-differentiation by computing the model's predictive accuracy on each teacher's students, controlling for gender imbalances in ability. Using the universe of French Grade 12 students' math transcripts, we show that math teachers use different words to assess the work of male and female students with a similar performance level. The analysis of our model's best gender predictors reveals that teachers insist more on positive managerial aspects for their female students, while equally performing males are both more criticized for their unruly behavior and praised for their intellectual skills. Using the within high school, elective and year variation, we further show that a higher gender differentiation in the teachers' feedback is associated with an improved performance at the math *baccalauréat* exam, but does not relate to students' matriculation in the following year.

*This paper is jointly written with Marion Monnet. It greatly benefited from discussions and helpful comments from Elliott Ash, Asma Benhenda, Anne Boring, Marc Gurgand, Sylvie Lambert, Dominique Meurs, Roland Rathelot, Michael Stepner, Clémentine Van Effenterre, and from all participants at the IPP Reading group, Sciences Po Education Policies 2020 seminar, Paris School of Economics Labor Chair Seminar 2021, ETH Zurich Data Science and Economics Seminar 2021.

Keywords: *higher education, teacher feedback, selective STEM*

Introduction

Stereotypes about different gender aptitudes in science-related fields are still pervasive and are now widely acknowledged as one of the causes of women's underrepresentation in Science, Technology, Engineering and Math (STEM) fields (Miller et al., 2014; Leslie et al., 2015). Stereotyping STEM as a male domain influences females' attitudes towards science, their performance in science subjects (Nosek et al., 2009a), and eventually their interest as well as their willingness to pursue a scientific career (Cheryan and Plaut, 2010).

The development of gender-related science beliefs starts at early ages (Bian et al., 2017), first shaped by parents' attitudes and beliefs towards science and, later on, by the schooling environment, namely teachers. Several studies show that math teachers hold stereotyped beliefs conveyed to their students through their classroom instructions (Keller, 2001), their own attitudes such as math-anxiety (Beilock et al., 2010), or through the type of feedback provided.¹ Teachers holding traditional gender views tend to adopt more person-based feedback insisting on fixed aptitudes (e.g. "You are gifted in math"). They convey the idea that math ability is innate and that females are less likely to have it (Copur-Gencturk et al., 2020), while more progressive teachers use effort-based feedback, with the idea that math ability can be acquired with efforts (e.g. "You have worked hard"). The type of feedback provided sends a signal about whether intelligence is fixed or malleable and is all the more crucial as females are more sensitive than males to the different feedback received. Effort-based feedback enhances females' motivation (Corpus and Lepper, 2007), their sense of belonging and willingness to pursue math (Good et al., 2012), while those aspects are harmed by person-based feedback.

This paper documents the extent to which the feedback used by math teachers is gender-differentiated, and how exposure to teachers using such a gender-differentiated feedback relates to Grade 12 students' academic performance and higher education enrollment choices.

To our knowledge, the analysis of teacher feedback has predominantly been done in experimental settings or using a limited sample of grade transcripts with a feedback analysis that remained mostly descriptive (Demoulin and Daniel, 2013). Our first contribution is to build a statistical model to shed light on teachers gender-differentiated feedback in a data-driven manner. We perform

¹See Alan et al. (2018); Carlana (2019a); Copur-Gencturk et al. (2020) for recent references documenting teachers' stereotyped beliefs.

a textual analysis of comprehensive and non-experimental data and propose a synthetic measure of teachers' gender-differentiated feedback based on the written feedback provided to students. Using text mining techniques on the universe of Grade 12 students' math transcripts over the period 2012–2017, we build a statistical model that predicts students' gender relying on the words used by math teachers. For each teacher, the share of correctly predicted observations, also known as the accuracy of the model, is our measure of teacher gender-differentiated vocabulary (GDV hereafter). The more a teacher makes use of female predictors to assess its female students' work and male predictors to assess its male students' work, the better the predictive accuracy, and the stronger the teacher gender differentiation. To get a sense of how different the vocabulary used by high GDV teachers is relative to gender-neutral ones, we further classify gender predictors into positive *versus* negative, and managerial *versus* competence-related categories.²

The extensive analysis of the vocabulary used in transcripts and the synthetic measure of gender-differentiated feedback proposed in this paper allow us to describe the feedback patterns used by French Grade 12 math teachers. We find that, on average, math teachers differentiate their vocabulary according to their students' gender, and document a large variation in the distribution of the teacher GDV measure. Another finding is that math teachers are significantly more differentiating their vocabulary than teachers in humanities. Using the feedback provided by teachers in those subjects and applying the same estimation procedure yields lower predictive accuracies, meaning that the words used by these teachers are more gender-neutral than those of math teachers.

The classification of the best gender predictors along the four above-mentioned dimensions reveals marked gender differences in the vocabulary used by math teachers to describe the work of equally able students. The first striking fact concerns the relative proportion of positive and negative feedback received by each gender. Out of the 30 best female predictors, two-thirds are classified as positive and mostly relating to their behavioral skills and efforts. The reverse is observed for male predictors, that relate to negative managerial aspects in two-thirds of cases. Positive male predictors however praise their intellectual skills. Overall, math teachers insist more on positive managerial aspects and encourage the efforts provided by their female students, while equally performing males are both more criticized for their unruly behavior and praised for their

²The words classified as “managerial” refer to the student’s attitude in class as well as the efforts and actions undertaken for the subject, while those classified as “competence” relate to math concepts and the school environment or to the students’ intellectual ability (Morgan, 2001).

intellectual skills.

The second contribution of this paper is to relate our teachers' GDV measure to students' academic performance and enrollment outcomes. We investigate the extent to which exposure to a high GDV teacher correlates with students' percentile rank at the national *baccalauréat* exam in different subjects. We also investigate the relationship with their higher education choices by looking at the rank-order lists that students submit when applying to higher education programs. Finally, we track students the year following their high school graduation and look at whether the students exposed to high GDV teachers make different enrollment decisions. Our identification strategy exploits the within high school variation in teacher GDV, and relies on the fact that teachers' allocation to the different classes is almost as good as random conditional on some observable characteristics.

We find that being assigned to a teacher with a one standard deviation higher GDV is associated with an increase in students' performance at the math *baccalauréat* national exam by 0.34 percentile rank on average (from a baseline of 50), the effect being slightly larger for female students (+0.37 percentile rank against +0.31 for males). Compared to those who are exposed to teachers from the bottom 10 percent of the teacher-GDV distribution, female students exposed to teachers from the 4th decile or above see their percentile rank in math increase by up to 2 percentile ranks. Most likely, teacher GDV is related to specific teacher attitudes and practices which could contribute to explain the better performance of their students at the national *baccalauréat* examinations. We find, however, that teachers with a higher GDV are not systematically higher quality teachers, as measured by their value-added, and are not more likely to overgrade female students relative to their male counterparts in continuous assessments. In addition, the coefficients obtained on math performance are robust to the inclusion of the teachers' grading bias measure as control, and are reduced but remain significant when including teacher quality as control.

Most of the effects on the type of programs students rank when applying to higher education or on the type of programs where they eventually matriculate in the following year are small in magnitude and not statistically significant. Being exposed to a one standard deviation higher teacher GDV leaves female students' likelihood of top-ranking a STEM program unchanged, and is associated with a slight decrease in that of male students, by 0.27 percentage point, significant at the five percent level. The effects on matriculation in the year following high school graduation are close to those on top-ranked programs.

Related literature. Our paper speaks to different strands of the literature. First, it relates to the broader literature on the origins of women’s underrepresentation in male-dominated fields. While this literature initially focused on gender differences in math performance, whereby males are said to outperform females in math test scores, a consensus seems to emerge on the fact that those test score gaps are quite limited (Lindberg et al., 2010), and are themselves influenced by social and cultural norms, as more egalitarian countries usually record lower gender performance gaps (Guiso et al., 2008; Breda et al., 2018). Cultural and social norms are now privileged as the most plausible explanation for such gender gaps. Social norms, as conveyed by parents, peers and teachers are major determinants of young females’ math perceptions, school performance and future educational choices.

Parents’ and teachers’ positive attitudes towards math tend to spill over to their children and pupils, who embrace a growth mindset, i.e., who believe that one can succeed by exerting effort (see Gunderson et al. (2012) for a review). Such a positive attitude is associated with a higher math performance and increased graduation in STEM fields (Cheng and Kopotic, 2017). Math teachers and parents working in STEM occupations may also act as role models, by breaking stereotypes and raising interest for science-related careers (Lim and Meer, 2017; Eble and Hu, 2019a; Cheng and Kopotic, 2017), even though they sometimes engage in different teaching practices or behavior with female and male students. For instance, Oguzoglu and Ozbeklik (2016) find that fathers in STEM occupations are much less likely to transmit STEM-specific tastes to their daughters when they also have a son, while Lavy and Sand (2018a) and Terrier (2020) find that math teachers’ grading behavior is biased in favor of female students. Peers also play a role in the perpetuation of gendered-beliefs and stereotypes. Classmates are a vector of their parents’ beliefs. Exposure to peers whose parents believe that males are better than females at learning math increases a child’s likelihood of believing it as well, harms females’ but improves males’ performance in math (Eble and Hu, 2019b). Peers may also convey such social norms by penalizing classmates who deviate from the norm (Skořajić et al., 2020).

Our paper also relates to the growing research in social sciences using text as data (Gentzkow et al., 2019) to uncover patterns of biases and discrimination. In an experimental setting where questions are randomly asked by a fictitious male or female account on a math forum, Borhen et al. (2018) show that the answers to questions posted by female accounts contain significantly more positive *and* negative words and opinions than answers to males. Wu (2018) analyses threads from

the Econ Job Rumor Forum and shows that the language used to assess women in academia is substantially different from that of males. When women are mentioned in a thread, the discourse is significantly less related to professional matters and focuses more on personal information or physical appearance. Koffi (2020) uses bibliometric data to investigate gender biases in citation patterns of Economics scholars. She shows that omitted papers, i.e., the ones that are relevant and should be cited, are 15 percent to 30 percent more likely to be female-authored than male-authored, and that this omission bias is twice as large in more theoretical fields involving mathematical economics. Ningrum et al. (2020) perform a textual analysis of job advertisement and find explicit discrimination against females in the hiring process.

The remainder of this paper is organized as follows. In Section 3.1, we provide some institutional background on the French secondary education system and on the admission procedure for higher education. In Section 3.2, we describe the different data sources that we use, along with some descriptive statistics on the population of Grade 12 science track students and their math teachers. Section 3.3 presents our empirical strategy. We detail the different steps to obtain our measure of teacher gender-differentiated vocabulary and our identification strategy to measure the impact of teacher GDV on students' outcomes. In Section 3.4, we first show some general descriptive statistics on the vocabulary used by math teachers, and provide a detailed analysis of the gender-differentiated vocabulary as well as some statistics on the distribution of our GDV measure. Section 3.5 investigates the relationship between teacher GDV and students' academic performance, preferences for higher education programs and enrollment outcomes in the year following high school graduation. Section 3.6 discusses the mechanisms potentially driving our results and Section 3.7 concludes.

3.1 Institutional Background

3.1.1 The French Secondary Education System

In France, the secondary education system consists of seven years of schooling, divided into four years common to all students and taught in middle schools (*collège*, Grade 6 to 9), and three years of high school (*lycée*, Grade 10 to 12), either delivering a vocational or a general and technological training. Both middle and high school curricula end with a national examination. At the end of

middle school, students take the *Diplôme National du Brevet* (DNB), which tests their knowledge and skills in math, French and history and geography. At the end of Grade 12, high school students take the *baccalauréat* examinations. Only students holding the *baccalauréat* can enter the higher education system.

In general and technological high schools, after a common *Seconde générale et technologique* year (Grade 10), students are tracked into a general (80 percent of students) or a technological curriculum (20 percent of students). General curriculum students further specialize by choosing their track (science, economics or humanities), when entering Grade 11, and their elective course, when entering Grade 12. Students tend to specialize according to both their comparative advantage and their preferences, which leads to marked gender patterns in track and elective course choices. While females are slightly underrepresented among science track students (they represented 47 percent of the science track students in 2018), the economics and humanities tracks are largely female-dominated: in 2018, 60 percent of economics track students and 80 percent of humanities track students were females (MENJ-MESRI 2019). These gender patterns in track choice are further reinforced by the choice of elective courses. The differences are particularly striking when focusing on science track students. Female students are largely overrepresented in the earth and life science elective, where they represent 63 percent of students, against only 30 percent in computer sciences and 15 percent in engineering. The proportions of female students in math and physics-chemistry electives are more balanced (42.5 percent and 48 percent respectively).

The gender segregation within French high schools is however limited beyond the segregation induced by the choice of an elective course. The composition of each class is determined by high school principals who, while taking into account the students' electives when defining the classes, also declare putting gender diversity on top of their priority list (Cnesco, 2015). Most principals also declare valuing some heterogeneity in terms of students' academic achievement level but, unlike for gender, the academic stratification within high schools remains substantial. Ly and Riegert (2015) have looked at the determinants of the within high school segregation and found that grouping students according to their elective courses accounts for two-thirds of the observed social and academic segregation.

3.1.2 Tertiary Application and Enrollment

High school students apply to post-secondary programs in the Spring term of Grade 12. Throughout the year, the head teacher guides students by providing assistance with the application procedure and some counseling regarding the choice of programs. At the end of the academic year, the high school principal gives an opinion on the students' chances of success that appears in the application files, but students remain free to apply to any program of their choice.

The post-secondary programs students can apply to fall into two broad categories, with, on the one hand, university programs, which are mostly non selective and open to all high school graduates, and, on the other hand, selective programs.³ The latter include three different types of programs, which have a strict academic stratification: two-year post-secondary vocational and technical programs (*sections de techniciens supérieurs* and *instituts universitaires de technologie*), post-secondary management and engineering schools, and the two-year elite *classes préparatoires aux grandes écoles* (CPGE). The CPGE prepare students to the entry exam to the most prestigious French colleges (the *grandes écoles*) in science, business, or humanities. Science CPGE are further specialized into three main categories: MP (math and physics), PC (physics and chemistry) and BCPST (biology), while business CPGE can be distinguished between sciences, economics and technological programs, and humanities CPGE are composed of classics and social sciences programs.

Until 2017, the post-secondary admission procedure was centralized through the *Admissions Post-Bac* (APB) online platform for most undergraduate programs. The main round of the student assignment mechanism relied on a procedure that was close to the college-proposing deferred acceptance algorithm (Gale and Shapley, 1962; Roth, 1982). Applicants were invited to submit a rank-order list of programs that could include up to 36 choices, with a maximum of 12 choices per type of program (University program, STS, CPGE, etc.). After the list's submission deadline, applicants were ranked by the different programs. For selective programs, the ranking was based on their Grade 11 and 12's academic records. The grades obtained in different subjects as well as teachers' written feedback played a crucial role in determining the programs' rankings of applicants. For non-selective programs, applicants were ordered according to some priority rules, based on their catchment area and the program's rank in the student's list.

³Since the 2018 reform of the application procedure that changed Admission Post-Bac into Parcoursup, universities are now allowed to select students according to their past academic performance but our study does not cover the post 2018 period.

3.2 Data and Summary Statistics

This section details the different data sources that we use to build our measure of teacher GDV and to quantify its impact on students' outcomes (Section 3.2.1). We also present summary statistics on the sample of Grade 12 students and on their math teachers (Section 3.2.2).

3.2.1 Data Sources

We use three main administrative databases: the college application data for six cohorts of Grade 12 students (2012-2017) as well as information on their math teachers collected via the APB platform; the higher education enrollment data; and the data of the two main national exams (*Diplôme national du brevet*, DNB, and *baccalauréat*).

APB data. Our primary source of information is the comprehensive application data from the APB platform over the period 2012–2017. A substantial amount of information is collected by this platform during the application process. We first use the students' digitalized academic records to retrieve teachers' feedback on all the subjects taken by students during Grade 11 (three trimesters) and Grade 12 (two trimesters only). This is the main input used to build our measure of teacher GDV. Teachers and students are uniquely identified in the data, which enables us to link these transcripts to students' and teachers' characteristics contained in a separate APB file. Along with basic sociodemographic information on the students (gender, place and date of birth, parental socio-economic status, etc...), the APB data provide detailed information on the schooling trajectories of the students in high school (school path, track and elective choices, etc...). The data also allow us to determine the teachers' gender, the subject they teach and whether they are the head teacher of the class.

Another interesting feature of these data is that they keep a record of the final rank-order list of programs submitted by each participating student. For each year, we also know the matching outcome, i.e., the program to which each student was admitted along with the students' acceptance decision (acceptance, conditional acceptance or rejection).

School performance data. We use the OCEAN database, managed by the French Ministry of Education, for the grades obtained in two national examinations: the *diplôme national du brevet*

(DNB), taken at the end of Grade 9, and the *baccalauréat*, taken at the end of Grade 12. We use the former to control for the students' past academic performance in the estimation procedure, while the latter is used as an outcome for student's performance at the end of high school. For that purpose, and to make grades comparable across years, we transform the initial grades ranging between 0 and 20 into percentile ranks, where 0 and 100 are respectively the ranks for the lowest and the highest performing students. It is worth mentioning that both exams are "blind" tests, i.e., they are anonymously and externally graded.

Enrollment data. To track Grade 12 students' enrollment outcomes in the following academic year, we use the *Système d'Information sur le Suivi de l'Étudiant* (SISE) for enrollment in non-selective post-secondary programs (*Licence*), which is managed by the Statistical Office of the French Ministry of Higher Education. This dataset, which covers the academic years 2012 to 2017, records all students enrolled in the French higher education system outside of CPGE, STS and post-secondary programs leading to paramedical and social care qualifications. For selective programs, we use a separate administrative data source called *Bases Post-Bac*. These comprehensive administrative registers cover the universe of students enrolled in selective post-secondary programs, i.e., CPGE and STS.

Sample restrictions. Given that the focus of this study is on the impact of math teachers' vocabulary, we restrict our sample to Grade 12 students enrolled in the science track as they are the ones interacting most frequently with their math teachers, relative to the humanities or the economics tracks.⁴ We then drop students for whom the math teacher's identifier or the grade transcript is missing, which represents about 50 percent of Grade 12 students from the science track in 2012 and goes down to 15 percent in 2017 (Table 3.1). In the vast majority of cases (between 70 and 95 percent of missing observations), teachers' identifiers and grade transcripts are missing because the entire high school is not reporting its students' grade automatically on the APB platform. Dropping those observations therefore amounts to dropping entire high schools and is not a threat for the internal validity of our analysis.⁵ We finally restrict our sample to

⁴The science track curriculum includes six hours of compulsory math classes (an extra two hours if the math elective is chosen) against four hours for the economics track (an additional hour and a half for the math elective) and none for the humanities track (four hours for the math elective).

⁵It might, however, affect the external validity of our analysis. Table 3.C1 in Appendix 3.C shows the OLS coefficients obtained when regressing a dummy indicating whether the high school has all grade transcripts missing on the high school's average characteristics. High schools with a higher share of female and free lunch students are more likely to be reporting the grade transcripts. Reassuringly, the relative performance of female vs. male students

high schools having at least two science track classes – as our identification strategy relies on a within-school comparison of students (see section 3.3), and to teachers having taught at least two classes over the period. These restrictions lead us to exclude between 6 and 20 percent of students. Once those restrictions have been applied and depending on the year considered, the sample of analysis consists of 40 to 75 percent of Grade 12 students from the science track, for a total of approximately 700,000 observations.

3.2.2 Summary Statistics

Students' characteristics. Table 3.2 provides summary statistics of Grade 12 science track students' characteristics for the whole sample of analysis and separately for male and female students. Students are eighteen years old on average and mostly come from a high (43 percent) or a medium high socio-economic background (16 percent).⁶ Female students are slightly underrepresented in the science track as they account for 47 percent of science track students but 54 percent of all general Grade 12 students (MENJS-MESRI, 2018). Turning to elective courses, we note striking gender differences. Half of female students opt for the earth and life science elective against only one fourth of males. Female students are also underrepresented in the math (19 percent against 27 percent) and engineering and computer sciences electives (6 percent against 20 percent). Another noticeable difference between male and female students relates to their past academic performance, as measured by the national percentile rank at the DNB math exam. Males' average rank is approximately 4 points above that of females. Figure 3.1 further shows that males are largely overrepresented in the top quartile of math performance: they make 58 percent of the top math achievers. On the other hand, females outperform males in French at both the DNB and *baccalauréat* exams, with an average percentile rank that is 10 points higher than that of their male peers. Both imbalances, in terms of elective choices and past school performance, are accounted for in our identification strategy and during the estimation procedure.

Students' higher education enrollment. The gender differences observed for elective courses are also found when looking at enrollment in the higher education system, as shown in Figure

at the math DNB examinations only marginally affects the probability of not reporting grade transcripts.

⁶Students' socioeconomic status (SES) is measured using the French Ministry of Education's official classification, which uses the occupation of the child's legal guardian to define four groups of SES: high (company managers, executives, liberal professions, engineers, intellectual occupations, arts professions), medium-high (technicians and associate professionals), medium-low (farmers, craft and trades workers, service and sales workers), and low (manual workers and persons without employment).

3.2. Both males and females enroll in a scientific program or in medicine in the same proportion: 54 percent of females and 58 percent of males opt for a STEM program (be it selective or not) or medicine. About one-fifth does not enroll in higher education and the remaining choose a non-scientific program (selective or not). However, we do observe substantial gender segregation within scientific subjects. The proportion of females enrolling in any STEM program is much lower than that of males. Only 11 percent of female students choose a selective STEM program against 21 percent for their male counterparts. Those proportions are respectively 16 percent and 26 percent for enrollment in a math, physics, chemistry or computer science program at university. Females rather enroll in a medicine program: 27 percent of Grade 12 females from the science track opt for such program against 11 percent of males only.

Math teachers' characteristics. Table 3.3 reports some descriptive statistics for the sample of math teachers in Grade 12 science track. There are 6,772 math teachers in the sample, 58 percent of which are males. A little more than half of them have been the head teacher of a class at least once over the period covered by our data. Those teachers are likely to have a stronger influence on students' performance and enrollment behavior as they counsel students on top of teaching them. Each teacher is in charge of only one Grade 12 class from the science track on average each year, with an average class-size of 28 students (90 percent of teachers teach only one Grade 12 science track class per year). Teachers appear almost four times in the sample, meaning that we have on average four teacher×classroom observations, which is crucial for the reliability of our GDV measure (see section 3.3). Finally, the average feedback is made of 7.4 words, with large variability in feedback length across teachers.

3.3 Empirical Strategy

The first part of this section explains the estimation procedure used to measure teachers' gender differentiated vocabulary (GDV) (Section 3.3.1). The second part presents the identification strategy to estimate the impact of teacher GDV on students' outcomes (Section 3.3.2).

3.3.1 Measuring Teachers' Gender Differentiated Vocabulary (GDV)

The measure of teacher GDV proposed in this paper leverages the rich data on teachers' feedback provided to students three times a year in their Grade 12 academic records. This personalized

feedback reflects the teachers' perception of the students' performance, work, and behavior in class throughout the year. To investigate whether the words used to characterize the students' work, behavior and ability differ by gender, we build a model that predicts students' gender based on the words used in the teachers' feedback. Using machine learning techniques, we estimate a model on the sample of Grade 12 science track students. We then use this fitted model to compute a measure of gender differentiated vocabulary for each teacher based on her observations only, controlling for class-level gender imbalances in students' prior academic performance. The different estimation steps are presented below while the detailed procedure can be found in Appendix 3.A.

Data preparation. Converting the corpus of teacher feedback into a statistical database is done in two steps. First, we rely on text mining techniques to replace each word by its root and hence ensure its gender neutrality (Gentzkow et al., 2019). Second, the corpus of teachers' feedback is turned into a matrix that contains one row per feedback and W_n columns, where W_n is the number of distinct words appearing in the corpus. Each of these columns is a dummy that takes the value one if the considered word appears in the student's feedback, and zero otherwise.⁷ Finally, we classify words into one of the four following categories, inspired from the sociology and psychology literature: positive (resp. negative) competence-related aspects and positive (resp. negative) managerial aspects. Ambiguous words (i.e., the ones used in both positive and negative contexts) or words that do not fit in any of the categories are respectively labeled neutral or unclassified.⁸ This classification is not directly used to build the measure of GDV but is necessary to characterize along which dimensions the vocabulary used by teachers differ.⁹

⁷Increasing the flexibility of the model by adding interactions between words (*bigrams*) does not improve the predictive accuracy of the model. We therefore stick to the simplest model specification, involving single words (*unigrams*) only.

⁸The classification of words is detailed in Tables 3.A2 and 3.A3 of Appendix 3.A

⁹We attempted to build these categories in a data-driven manner using bi-term topic models tailored for short texts, but these models performed poorly on our data. Our data is indeed quite specific in that texts are very short, with an average number of tokens equal to 7, the overall vocabulary is quite limited ($\simeq 1,600$ words) with little variation in the topics used as they all relate to academic performance and behavior. We therefore faced the typical challenges inherent to such short texts: the generated topics gathered inconsistent words (*trivial topics*) and the different topics were highly similar with a lot of words in common (*repetitive topics*, see Wu et al. (2020) for a discussion on those issues.)

Student gender estimation and prediction. We assume that the probability of being a female student conditional on the words used in the feedback has a logistic form:

$$P(Female_i = 1|W_i) = \frac{\exp(\alpha W_i)}{1 + \exp(\alpha W_i)} \quad \forall i \quad (3.1)$$

Our objective is to find the set of α coefficients that maximize a penalized version of the log likelihood $\ln(L(\alpha))$ of observing the realized gender distribution, where λ is the regularization parameter:¹⁰

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} (-\ln(L(\alpha)) + \lambda \sum_{w=1}^{W_n} |\alpha_w|) \quad (3.2)$$

The model described by Equation (3.1) is trained on a subsample of Grade 12 students. Using the set of $\hat{\alpha}$ coefficients retrieved from the estimation procedure, we use the hold-out sample to predict each student's gender as follows:

$$\hat{P}(Female_i = 1|W_i) = \frac{\exp(\hat{\alpha} W_i)}{1 + \exp(\hat{\alpha} W_i)} \quad \forall i \quad (3.3)$$

On average, our model correctly predicts students' gender in 60 percent of cases, and performs better at predicting males' gender than females', as 67 percent of males are correctly classified against 53 percent for females.

Teachers' gender differentiated vocabulary (GDV). We define each teacher j 's GDV for class c as the predictive accuracy of the model described by Equation (3.1) fitted on his or her students only, where the predictive accuracy is the share of the teacher's students whose gender is correctly predicted by the model.¹¹

Since gender is correlated with math performance (see Section 3.2.1), the model described by Equation (3.1) is likely to perform better on classes with stronger gender imbalances in terms of math performance. To alleviate this concern, we predict teacher j 's GDV for class c on a balanced subsample of students. More specifically, for each teacher, we undersample as many males and females from each quartile of prior math ability, so that we have 50 percent of male and female

¹⁰We select the regularization parameter λ using a 10-fold cross-validation procedure, and pick the highest λ value such that the error of prediction lies within one standard deviation of the minimal error (Hastie et al., 2009).

¹¹A student is classified as female if her predicted probability of being a female is larger than 0.5, otherwise she is classified as male.

students at each ability level.¹² The quality of the prediction could also be influenced by the feedback’s length which varies by teacher (Table 3.3). For lengthy feedback, defined as the ones with an above-median length, we randomly sample the median number of words.

Another related concern is that teacher j ’s GDV estimated for class c could capture some unobserved class-specific gender differences in behavior or performance. To rule out this possibility, we compute an alternative measure that we call the *leave-one-out* GDV, defined as the average of teacher j ’s GDV over all the classes she taught during the study period, excluding class c . Our two measures are defined as follows:

$$GDV_{jc} = \frac{1}{N_{jc}} \sum_{i=1}^{N_{jc}} \mathbb{1}\{Sex_i = \widehat{Sex}_i\} \times 100 \quad \forall j, c \quad (3.4)$$

where N_{jc} is the number of students in the balanced subsample of teacher j ’s students from class c , and:

$$GDV_{j \setminus c} = \frac{1}{N_j - 1} \sum_{c' \neq c} GDV_{jc'} \quad \forall j, c \quad (3.5)$$

where N_j is the number of classes that teacher j taught throughout the period under study.¹³

Both GDV measures theoretically lie between 0 (the model systematically misclassifies females as males and males as females) and 100 (all students are assigned their observed gender). The higher the accuracy for a given teacher, the better we can, on average, recover his or her students’ gender based on the words she uses in her feedback, and hence the stronger the gender differentiation in the vocabulary that she uses in her assessments. A model that randomly assigns each student a gender with probability 0.5 would achieve an accuracy of 50 percent, meaning that our model predicts gender better than random guessing for all teachers whose accuracy is above 50 percent. An accuracy below 50 percent is possible in our setting given that the prediction is done on small samples at the teacher level. However accuracies are averaged over 100 estimations to limit such random fluctuations, and the *leave-one-out* GDV is itself an average of several accuracies, therefore reducing the noise inherent to the measure.

¹²Prior ability in math is proxied by the percentile rank in math at the DNB exam.

¹³In practice, both GDV measures are computed as the average of GDV and *leave-one-out* GDV estimated on 100 random balanced subsamples of teacher j ’s students.

3.3.2 Identification Strategy

The second objective of this paper is to establish a causal link between the teacher GDV measure presented above and students' performance and enrollment outcomes. Our identification strategy relies on the comparison of students enrolled in the same high school but exposed to math teachers with different levels of GDV. More specifically, we exploit the within high school variation in teacher GDV and estimate the following equation, controlling for year and elective courses fixed effects:

$$Y_{isjet} = \alpha + \beta GDV_{j\setminus c} + \gamma_s + \eta_e + \delta_t + \epsilon_{isjet} \quad (3.6)$$

where Y_{isjet} is the outcome of student i in high school s with elective courses e taught by teacher j during academic year t . $GDV_{j\setminus c}$ is teacher j 's standardized GDV measure and is class-specific as we use the *leave-one-out* GDV described in Equation (3.5). γ_s is a set of high school dummies and η_e and δ_t are elective and year fixed effects respectively. The coefficient of interest is β , which measures how a student's outcome is affected by being assigned a teacher with a one standard deviation higher GDV. The standard errors are robust and clustered at the teacher level.¹⁴

The validity of our identification strategy requires that teacher GDV is not systematically correlated with students' characteristics. We formally test this in Section 3.5.

3.4 Math Gender-Differentiated Feedback

In this section, we first provide an overall picture of the gender differentiation of the vocabulary used by math teachers (Section 3.4.1), before turning to the analysis of our GDV measure (Section 3.4.2).

3.4.1 Descriptive Statistics on Math Feedback by Gender

This section aims at providing a comprehensive picture of what a math feedback looks like for a Grade 12 male or female student. We first provide statistics on the distribution of word counts of math feedback overall and by type of feedback, and then provide statistics aggregated at the teacher level.

¹⁴Although standard errors should be bootstrapped to account for prediction error, we do not implement this correction due to computational limitations.

Students' math feedback. Panel (a) of Figure 3.3 displays basic summary statistics on the distribution of the number of content words appearing in the math feedback received by male and female students separately.¹⁵ The first finding is that female students tend to receive slightly shorter feedback: while the median feedback addressed to male students is made of seven content words, the median female student feedback contains only six content words. The upper tails of the feedback length distributions are similar across gender, with 25 percent of feedback containing more than nine content words, but the lower tails differ slightly, with 25 percent of feedback containing less than five words for male students, and less than four words for female students.

Those summary statistics are then broken down according to the four dimensions mentioned in Section 3.3: managerial vs. competence-related feedback and positive vs. negative feedback. The summary statistics along the positive vs. negative dimensions do not reveal significant differences between male and female students. For both genders, the median feedback contains two positive words and zero negative word, and 25 percent of the students receive more than three positive words and one negative word. The managerial and competence-related dimensions highlight stronger gender patterns at the top of the distribution only. The median feedback addressed to male and female students contains one competence-related word and two managerial-related words. However, while 10 percent of male students receive more than four competence words, this is the case of only eight percent of female students.

Math teachers' feedback. In Figure 3.4, we take a teacher perspective to dig into the gender differentiation of the type of feedback provided. For each teacher, we compute the share of managerial or competence-related feedback separately by gender, and plot the distributions of teachers' gender gaps in the shares of managerial and competence-related feedback in Panel (a).¹⁶ While 81 percent of the teachers put a stronger emphasis on managerial matters in the feedback addressed to females compared to males, only 18 percent of the teachers insist more on competence-related aspects when they assess a female student's work and performance. We further characterize these managerial and competence-related feedback along the positive and negative dimensions. Panel (b) shows the differential propensity to address a positive feedback across gender: 80 percent of the teachers tend to be more positive for females than for males on managerial-related issues, while this is the case of only 20 percent of the teachers for competence-related matters. The propensity to be negative

¹⁵For simplicity, we define as content words all words that are not stop words (cf Section 3.A).

¹⁶The shares are computed on the subset of words that could be classified either as competence-related or managerial.

on managerial-related feedback is more balanced (44 percent of the teachers are more negative with females), while the weight given to negative feedback on competence-related matters is more pronounced for females, with 77 percent of the teachers that are relatively more negative with them.

3.4.2 Math Teachers' Gender Differentiated Vocabulary (GDV)

The naive analysis performed in the previous section reveals clear gender patterns in math teachers' feedback, but the observed differences are potentially resulting from genuine gender differences in student ability and attitudes towards math. We now turn to the analysis of our GDV measure estimated with the methodology described in Section 3.3, which is explicitly designed to isolate feedback patterns that are independent from students' own characteristics.

Distribution of teachers' GDV. Figure 3.5 shows the density as well as the cumulative distribution of the GDV and *leave-one-out* GDV measures separately. It provides evidence of the existence of a correlation between students' gender and the feedback received in math, controlling for the students' prior ability level in math. Our model predicts gender better than random for 90 percent of math teachers if we consider the GDV measure and for over 95 percent with the *leave-one-out* GDV.¹⁷ It correctly predicts the gender of 61 percent of the students for the median teacher.¹⁸ When breaking down the GDV distributions by teachers' gender, we note that, on average, female math teachers differentiate their vocabulary slightly more than their male colleagues (see Figure 3.6). Another finding is that using a gender-specific vocabulary is something quite specific to math-intensive subjects. We replicate the teacher GDV estimation procedure on the feedback provided to our sample of Grade 12 students in the following core subjects: physics & chemistry, biology, philosophy and modern language 1 and 2. Figure 3.7 shows that the *leave-one-out* GDV distribution for humanities-related subjects is shifted to the left compared to science-related subjects. This suggests that teachers in philosophy and modern languages are, on average, less inclined to use a gender-specific vocabulary in their feedback compared to math, physics and chemistry teachers, while biology teachers are somewhere in-between. Philosophy is a particularly relevant point of

¹⁷In total, only 4 percent of teachers have a *leave-one-out* GDV below 50 percent, the vast majority of which falls between 44 percent and 50 percent. This is mostly explained by the fact that those teachers are observed 3 times on average, compared to 4 times for other teachers. Their *leave-one-out* GDV is therefore slightly noisier.

¹⁸For comparison, when predicting whether the student's performance at the math DNB examinations corresponds to the top or bottom half performance in her class, the model achieves a median predictive accuracy of 65.8 percent.

comparison as its teacher-gender composition is close to that of math-intensive subjects (62% of males, see Table 3.B). Yet, philosophy teachers seem to use a more-gender neutral vocabulary.

Qualitative analysis of the best predictors. A high degree of gender differentiation in the vocabulary used, as measured by a high GDV measure, can reflect different teachers' attitudes. A gender differentiated feedback may be the expression of gender stereotypes regarding students' ability in math, but may also reveal a teacher's effort to adapt her feedback to the different student profiles she perceives. In order to better understand the extent to which the gender differentiation of the vocabulary used by math teachers reflects either of the two attitudes, we analyze the best gender predictors and classify them using the four categories defined in Section 3.3 to ease the comparison between male and female students' feedback.

The analysis of the best predictors of each gender points towards marked differences in the words used by teachers. Figure 3.8 reports the odds ratios derived from the estimation of the model described by Equation (3.1), for the top 10 predictors of each gender. A feedback mentioning the student's confidence level, her propensity of getting discouraged or her cheerful aspect ("smiles") is between 1.8 and 2.3 times more likely to be directed to a female than to a male student, relative to a feedback that does not mention it. Teachers are also more likely to mention that female students are stressed or panicked, and to insist on their exemplary conduct ("exemplary", "studious"). On the other hand, a feedback describing the student as childish ("childish", "has fun"), insisting on the need for careful handwriting, or praising the student's curiosity and intuitions is about 2.5 times more likely to be that of a male rather than that of a female student, compared to other feedback.

Figure 3.9 extends the analysis to the 30 best predictors of each gender and plots them on a quadrant that distinguishes positive from negative words (neutral words being in the middle), and where the colors refer to competence, managerial or unclassified words. The first striking feature of this graph is the relative proportions of positive versus negative type of feedback by gender. Among the top 30 male predictors, only five correspond to a positive feedback, while roughly two thirds of the best female predictors can be considered as positive. Most interestingly, conditional on being positive, the best male predictors almost all assess the student's competence-related aspects ("curious", "idea", "interest", "intuition"), while nearly all of the best female predictors assess managerial aspects ("irreproachable", "willingness", "persevere"). On the other hand, more than 80 percent of the best male predictors can be classified as negative and the vast majority refer to a

disruptive behavior (“has fun”, “childish”) or to a neglected work-effort (“waste”, “superficial”).

Finally, these results are confirmed when we consider all gender predictors. Figure 3.10 first shows the proportions of negative, positive and neutral feedback conditional on having a competence-related feedback (left-hand side of the graph) or a managerial one (right-hand-side). Among the predictors for being a female student that can be classified as competence-related, only 20 percent correspond to a positive feedback against 39 percent for male predictors, while 17 percent versus 12 percent are negative, the rest being neutral. Symmetrically, among the female predictors that can be classified as managerial, 44 percent correspond to a positive feedback against 27 percent for males. The latter get a much larger share of negative feedback: as much as 45 percent of managerial male predictors are negative while this proportion is only 29 percent for females.

Turning to the proportions of competence, managerial and neutral words conditional on having a positive or a negative feedback, we see that conditional on being positive, top female predictors assess their competence-related skills in only 16 percent of cases against 40 percent for their male counterparts. Regarding the breakdown of negative predictors, 40 percent of negative male predictors relate to managerial matters against 33 percent for female predictors, and those proportions are respectively 16 percent and 9 percent for negative competence-related predictors.

Vocabulary used by teachers' decile of GDV. The last thing we investigate is whether teachers with varying degrees of GDV differ from each other, by comparing the teachers' gender gaps in the share of positive words among competence and managerial-related feedback by decile of GDV (see Figure 3.11). Panel (a) plots the absolute values of the teachers' gender gaps and Panel (b) displays the share of teachers having a gender gap in favor of females, separately for competence and managerial-related feedback. The gender gaps in the share of positive words increase at a growing pace with GDV deciles. This is true for both managerial-related and competence-related feedback, with a gender gap that goes from 6 to 7 percentage points in the lower GDV deciles up to 10 percentage points in the 10th decile. In line with the findings resulting from the analysis of the gender predictors, Panel (b) reveals that higher GDV teachers are overwhelmingly relatively more positive with females on managerial-related feedback, and more negative with females on competence-related feedback.

Taken together, these descriptive statistics indicate that teachers do use a differentiated vocabulary for their male and female students. They seem to insist more on positive managerial aspects

and to encourage the efforts provided by their female students, while equally performing males are both more criticized for their unruly behavior and praised for their intellectual skills. In the following section, we investigate to what extent this gender differentiation in the feedback provided affects students' performance and future enrollment outcomes.

3.5 Impact of Teachers Gender-Differentiated Vocabulary on Students' Outcomes

After having documented differences in Grade 12 math teachers' gendered vocabulary, we turn to the impact of teachers' GDV on students' outcomes. We first perform a series of statistical tests aimed at validating our empirical strategy (Section 3.5.1). We then discuss the impact of teachers' GDV on students' academic performance and on their higher education choices and enrollment the year following high school graduation (Section 3.5.2). We show that our results are robust to a series of alternative specifications.

3.5.1 Validity of the Empirical Strategy

3.5.1.1 Exogeneity Assumption

The validity of our identification strategy requires that teacher GDV is not systematically correlated with students' characteristics. Ideally, we would want teachers to be randomly allocated to classes within a high school for a given elective course. We formally test this below.

Balancing tests between students' characteristics and teacher GDV. Table 3.4 reports the coefficients from a regression of the teachers' standardized *leave-one-out* GDV, defined at the class level, on students' socio-economic characteristics and baseline academic performance, along with a set of high school, year and elective courses fixed effects. The table shows that teacher GDV is not systematically correlated with students' observable characteristics. Out of the twelve characteristics included in the regression, only the age variable is significant at the 10 percent level, and the magnitude of the coefficient is very low.¹⁹ When jointly tested, we cannot reject the null hypothesis that baseline variables are not correlated with the *leave-one-out* GDV measure, thus

¹⁹The coefficient can be interpreted as follows: Increasing students' age by one year is associated with a decrease in teacher GDV by 0.0043 standard deviation.

providing evidence of teachers' random allocation conditional on high school, elective and year fixed-effects.

Random allocation of students. To check that students are randomly allocated to teachers within a given high school, elective course and for a given year, we follow Lavy and Sand (2018a) by performing a series of Pearson's Chi-square tests of independence. For each of the 27,668 high schools, elective course, and year unique combinations, we tabulate math teachers' identifiers with each of the students' baseline characteristics and test for independence.²⁰ Table 3.5 reports the percentage of p -values below the nominal values of 0.05 and 0.01. Except for the female dummy, we find that the empirical p -values are close to the nominal values (between 4.5 percent and 8 percent of p -values are below nominal levels, i.e. significant). For the female dummy, the empirical p -value is 11 percent, suggesting that in 11 percent of high school \times elective \times year combinations, we cannot exclude the non-random assignment of female students to classes at the 95 percent-level.

Taken together, the tests performed in this section suggest that in a given high school, elective course and for a given year, students are close to being randomly allocated to classes. To ensure that the results presented in Section 3.5.2 are not driven by the slight gender imbalances, Equation (3.6) is also estimated with the average proportion of females in the class as an additional control, as well as the full set of students' baseline characteristics.

3.5.1.2 Reverse Causality

Another concern regarding the GDV measure is that teachers' behavior could be influenced by the type of students' they are exposed to. In this case, our measure would not pick up some stable trait in the teachers' gendered vocabulary. This type of reverse causality is unlikely to be an issue in our setting.

First, we have shown in Table 3.4 that students' observable characteristics are rather well balanced across the distribution of teacher GDV: teachers that are more or less differentiating their vocabulary are not systematically assigned a specific type of students.

Second, each class is assigned its teacher's *leave-one-out* GDV measure, i.e., the average GDV measured in all the classes ever taught by the teacher except the considered class, which ensures

²⁰Continuous baseline characteristics such as age are previously dichotomized. The newly created variables take the value 1 if the student is above the median value and 0 otherwise. Measures of academic performance such as the students' percentile rank in DNB examination are transformed into quartiles.

that students do not contribute to the GDV measure they are being assigned.

Third, looking at the distributions of leave-one-out GDV measures estimated for other subjects further highlights the specific nature of science-related subjects, for which students' gender is better predicted on average than for humanities-related subjects (see Figure 3.7). Students' gender is correctly predicted in 59 percent of cases for humanities-related subjects, against 61 percent for math or physics and chemistry on average. Science-teachers leave-one-out GDV distributions are shifted to the right compared to that of humanities-related subjects. This suggests that on average, for a given class, science teachers differentiate more their vocabulary by gender. Our measure therefore captures differences that go beyond class-specific characteristics.

Finally, the fact that teachers' GDV is computed for multiple years and classes offers the opportunity to measure the persistence of teachers' GDV across classes and years, to ensure that we are not only capturing noise. The correlation between the teacher's GDV and leave-one-out GDV, i.e., the correlation between a given GDV and its average computed in other years \times classes, is 0.184 and is statistically significantly different from zero.²¹ It is worth noting that as we are correlating several GDV measured with error because of the small sample size used to make the prediction at the class level, this correlation suffers from an attenuation bias. As a comparison, in the teacher value-added literature, the within-teacher correlation is usually around 0.3 (Chetty et al., 2014).

Overall, we are confident that our GDV measure captures some persistence in the teachers' behavior.

3.5.2 Impact of Teacher GDV on Performance and Enrollment

Results on Academic Performance. Panel (a) of Table 3.6 reports the estimated impact of teacher's *leave-one-out* GDV on students' percentile rank at the math *baccalauréat* exam, obtained from the estimation of Equation (3.6). As a placebo test, the table also displays the teacher GDV's impact on the students' percentile rank at the philosophy *baccalauréat* exam.

A one-standard deviation increase in the math teacher GDV is associated with an increase in math performance at *baccalauréat* by 0.34 percentile rank on average, significant at the one percent level. In other words, this is the effect of moving from an average teacher in terms of GDV to a teacher from the 86th percentile of the GDV distribution. The effect is slightly larger for females, whose percentile rank increases by 0.37 when exposed to a one standard deviation higher GDV

²¹This correlation is obtained by regressing the teacher GDV on his leave-one-out GDV. The significance we refer to in the text tests for whether the regression coefficient is statistically different from zero.

teacher, against 0.31 for males, but the difference is not statistically significant. As expected, the math teacher's GDV has no effect on the percentile rank in philosophy, which is evidence that the GDV measure affects students' outcomes through the considered teacher only, and does not capture class-specific effects. Those results hold when we control for students' baseline characteristics, for the share of females in the class or for the average GDV measured in other subjects for students from the same class (see Table 3.D1 in the Appendix).

Even though moderate, these average effects hide heterogeneous responses that depend on the degree of gender differentiation of the teacher's vocabulary. Instead of including the teacher GDV linearly in the equation, we explore the intensity of the treatment by regressing the students' outcomes on a set of GDV deciles. The first (last) decile corresponds to the bottom (top) 10 percent of the math teachers' leave-one-out GDV distribution. Figure 3.12 plots the deciles' coefficients along with their 95 percent confidence intervals, separately for males and females. The figure shows that the higher the teacher GDV, the higher the effect on math performance, and even more so for female students. Compared to females exposed to the bottom 10 percent of teachers in terms of GDV, those exposed to teachers from the 4th decile or above see their percentile rank at *baccalauréat* in math increase by 1.5 to 2 percentile ranks on average, significant at the 5 percent level. The effect for females whose math teacher's GDV falls below the 4th decile ranges between 0.7 and 1.5 and is also statistically significant at the five percent level. This trend is also observed for males, but to a lower extent. Deciles' coefficients range between 0 and 1.5 percentile ranks, and are always below that of their female counterparts.

We also explored various dimensions of heterogeneity such as students' prior math performance or socio-economic status but did not find evidence of differentiated effects. Results are reported in Appendix 3.D.

Results on choice lists and matriculation in the following year. We then look at the effects associated with a one standard deviation increase in teacher GDV on the type of top-ranked STEM programs in students' final rank-order lists (Figure 3.13 and Table 3.6, Panel B), and on their actual matriculation in the year following high school graduation (Figure 3.14 and Table 3.6, Panel C). Being exposed to a teacher with a one standard deviation higher GDV does not alter females and males' STEM first choice for higher education programs nor does it affect their enrollment outcomes in the year following high school graduation. If anything, males are only marginally less likely to enroll in a selective program in the following year (-0.22 percentage

point) or in a vocational program (-0.10 percentage point), which represents a one percent and a 2.25 percent decrease with respect to the baseline proportions of males enrolling in such programs, as reported in Figure 3.2. We do not document any heterogeneous effects by deciles of teacher GDV, by student's prior math performance or by student's socio-economic status.

3.6 Mechanisms

In this section, we explore the mechanisms that could potentially drive the effects of math teachers' GDV on math performance at *baccalauréat*. We first look at whether teachers using a gender differentiated vocabulary are also encouraging females by overgrading them relative to males (Section 3.6.1). We then compute a measure of teacher quality to investigate whether math teachers with a higher GDV are also better teachers (Section 3.6.2). Finally, we investigate whether the male-specific or female-specific vocabulary has a different impact on math performance (Section 3.6.3).

3.6.1 Teacher Grading Bias

One first way through which teachers could encourage females and increase their performance is by overgrading them relative to their male peers. This teacher grading bias in favor of female students and its positive impact on school performance and enrollment choices has already been documented by Lavy and Sand (2018a) and Terrier (2020). We use their methodology and estimate teachers' grading bias by taking the difference between the gender gap in math test scores at the Grade 12 continuous assessment and this gender gap at the math *baccalauréat* exam (see Appendix 3.E for details). A negative (positive) grading bias is indicative of a grading bias in favor of females (males) at the continuous assessment. Consistent with what was previously found in the literature, we find that, on average, high school math teachers have a grading bias in favor of females (Table 3.E1).

Turning to the correlation between the grading bias and our measure of teacher GDV, we see from Figure 3.15 Panel (a) that a one standard deviation increase in teacher GDV is associated with a -0.07 standard deviation decrease in the grading bias, significant at the one percent level. This correlation is moderate but suggests that teachers who differentiate their vocabulary more are also slightly more likely to encourage females through higher continuous assessment grades. However, controlling for the grading bias in the main specification does not affect the magnitude

of the teacher GDV effect. Table 3.7 (Columns 1 and 3) reports the coefficients estimated on teacher GDV and on the teachers' grading bias for the main outcomes of interest.²² Except for the percentile rank at the math *baccalauréat*, the teacher grading bias almost never significantly affects males' or females' outcomes and, in any case, its inclusion as a control does not change the coefficients on teacher GDV. We can therefore exclude teachers' grading bias as a mediator of the impact of teacher GDV.

3.6.2 Teacher Quality

We now investigate whether the positive effects resulting from exposure to a higher GDV teacher can go through teacher quality.

We compute a measure of teacher value-added following the three steps described in the Chetty et al. (2014) paper. We first regress the percentile rank at the math *baccalauréat* exam on a set of students' baseline characteristics, variables accounting for students' past performance, and teachers' fixed effects. We predict residuals and use those students' residualized test scores to compute the average residualized test scores for each class×year combination. Class residuals in year t are regressed on their lags and leads, whose coefficients are the shrinkage factors. The coefficients obtained are finally used to predict teachers' value-added in year t . All the details of the estimation as well as the distribution of teacher value-added can be found in Appendix 3.E.

We then check whether teachers' GDV is correlated with their quality. Panel (b) of Figure 3.15 suggests a small yet significant quadratic relationship, where teachers with GDV measures that are two standard deviations below or above average have a slightly lower value-added. We control for teacher quality in Equation (3.6) and see that the impact of teacher GDV on percentile rank in math is slightly reduced with this additional control, going from +0.31 percentile rank to +0.25 percentile rank for males and from 0.37 percentile rank to 0.28 percentile rank for females (Table 3.7, columns 2 and 4.). Even though reduced, the impact of teacher GDV on students' math performance remains significant. This therefore suggests that our results are only partially channeled through teachers' quality.

²²See Table 3.E2 in Appendix 3.E for the teacher GDV coefficients on all the outcomes of interest when controlling for teacher grading bias or value-added.

3.6.3 Teacher's Type of Vocabulary Used

The last mechanism we explore is whether the type vocabulary used by teachers, i.e., the male or female-specific vocabulary, triggers different responses from students. For that purpose, we split the teacher GDV measure into two submeasures. For each teacher \times class, we compute the share of correctly predicted females on the one hand (hereafter referred to as the GDV-females measure) and the share of correctly predicted males on the other hand (GDV-males measure). These two measures enable us to highlight different patterns in the vocabulary used by teachers:

1. GDV-females = GDV-males: a teacher for whom both measures are very close has an overall GDV that is equally due to the correct classification of males and females by the model;
2. GDV-females > GDV-males: in this case, the overall teacher GDV is predominantly explained by the fact that the teacher uses the female-specific vocabulary with its female students and therefore our model better predicts females over males for that teacher.
3. GDV-females < GDV-males: this is the reverse situation where the overall teacher GDV is predominantly explained by the use of the male-specific vocabulary with male students, entailing a better prediction of males' gender over females'.

Panel (a) of Figure 3.E2 displays the distributions of the overall *leave-one-out* GDV and of the *leave-one-out* GDVs computed on males and females separately.²³ This graph shows that males are more often correctly classified by our model given that, on average, 66 percent of males' gender is correctly predicted against 55 percent for females. Panel (b) plots the correlation between both GDV submeasures and further shows that teachers for whom one gender is often correctly predicted have a substantially lower proportion of correctly classified observations for the other gender. A one standard deviation increase in the teacher GDV-males measure is associated with a 0.7 standard deviation decrease in the GDV-females measure.

To measure the extent to which the positive effect on math performance at *baccalauréat* is due to teachers using mostly the vocabulary associated to females or that associated to males, we estimate Equation (3.6) replacing teacher GDV by both GDV-males and GDV-females submeasures. Table 3.8 reports the estimations results for the main outcomes, and shows that the positive effect on the percentile rank at the math *baccalauréat* documented for higher-GDV teachers is primarily driven by teachers with a higher GDV-females, i.e., teachers who are more inclined to use the

²³Note that the density curve for the overall *leave-one-out* GDV is the same as the one displayed in Figure 3.5.

vocabulary associated to females. A one standard deviation increase in teachers' GDV-females (GDV-males) is associated with a 0.49 (0.39) percentile rank increase at the math *baccalauréat* exam on average, controlling for the average GDV-males (GDV-females). The positive effect of being exposed to a higher GDV-females teacher is stronger for female students, for whom the percentile rank increases by 0.52 against 0.34 when exposed to a higher GDV-males teacher, while these estimates are respectively 0.48 and 0.45 for males. These results suggest that the vocabulary used by higher GDV-females teachers is more prone to raising female students' performance and, to a lower extent, that of males. These results should however be interpreted with caution, as differences between male and female students are not statistically significant.

The magnitude of coefficients on the probability of top-ranking a STEM program or on matriculation outcomes are too low to substantially change students' program ranking or matriculation in the following year.

3.7 Conclusion

Relying on text mining techniques, we explore the vocabulary used by Grade 12 math teachers to assess their male and female students' work and propose a synthetic measure of the gender differentiation in teachers' vocabulary. This measure is designed to isolate feedback patterns that are independent from students' characteristics, and in particular from their ability level.

Using comprehensive administrative data on the universe of Grade 12 students' transcripts, we predict students' gender based on the feedback vocabulary, and compute the proportion of students whose gender is correctly predicted by our model separately for each teacher. This is our measure of teacher gender-differentiated vocabulary (GDV). We provide evidence that, on average, math teachers differentiate their vocabulary based on gender, and even more so when the teacher is a female. While this gendered-vocabulary exists in every core subjects, we find that teachers in science-related courses differentiate their vocabulary even more than teachers in humanities-related subjects. The qualitative analysis of the best gender predictors reveals that teachers insist more on positive managerial aspects and encourage the efforts provided by their female students, while equally performing males are both more criticized for their unruly behavior and praised for their intellectual skills. This gender differentiation is stronger as we move up in the teacher GDV distribution.

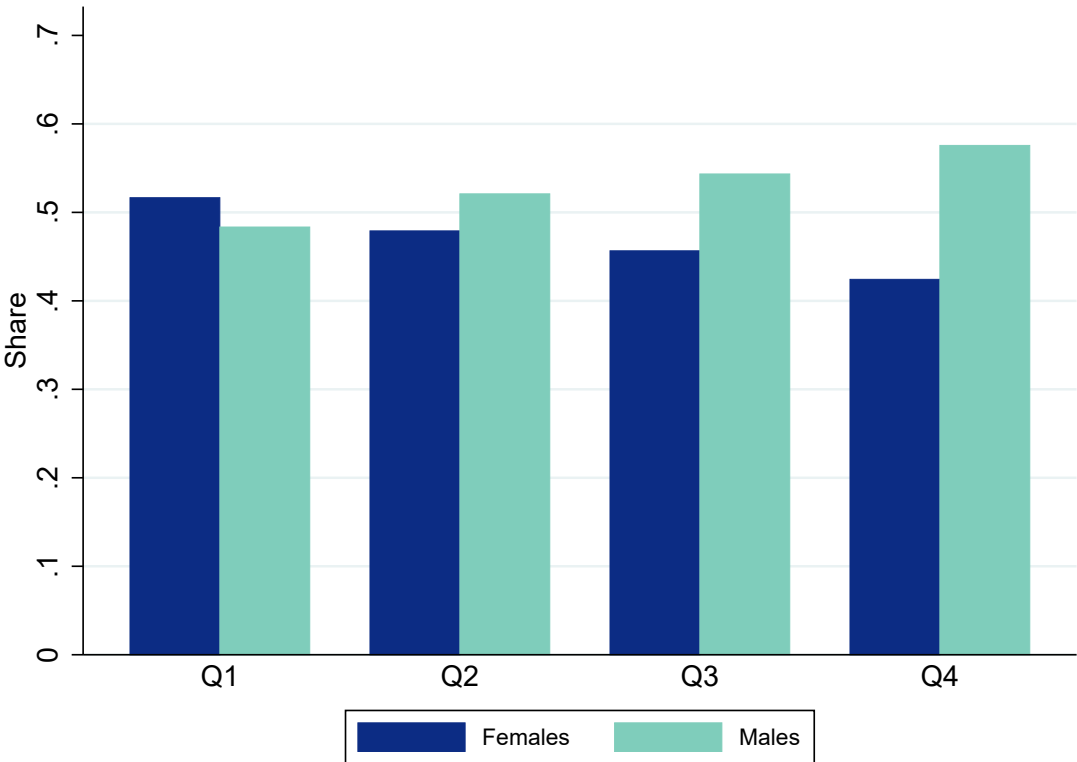
Exploiting the fact that, conditional on elective courses, the assignment of teachers within high

school is quasi-random, we establish a causal link between teachers' GDV and related teaching attitudes and students academic outcomes. Being exposed to a teacher with a one standard deviation higher GDV is associated with an increase in math performance at *baccalauréat* by 0.34 percentile rank on average, with slightly larger effects for female students. This effect is larger for students exposed to teachers with above-median GDV: relative to students exposed to a teacher from the bottom 10 percent of the GDV distribution, those exposed to above-median teachers see their percentile rank increase by up to 2 for females. A similar trend is observed for males, but to a lower extent. Deciles' coefficients range between 0 and 1.5 and are always below that of their female peers. We do not find any significant impact of teacher GDV on students' rank-order lists nor on students' matriculation in the following year.

Finally, we explore potential mechanisms driving the effects on math performance. We rule out the possibility that our results are driven by other teacher gendered behavior such as grading bias, but find that our results are partially channeled through teacher quality. We further show that our results are driven by teachers using the vocabulary associated to female students, i.e., teachers underlining the positive behavior and efforts, compared to teachers using the vocabulary associated to male students.

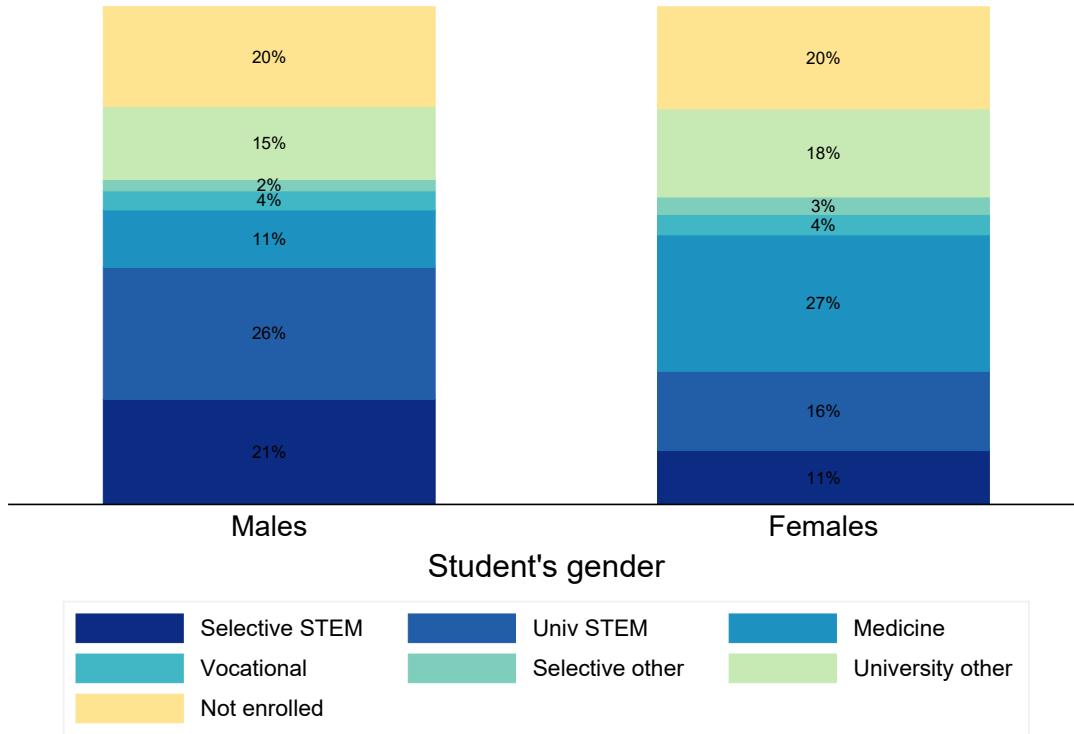
Our paper is at the crossroad between several social sciences, which offers a range of perspectives for future research. One pending question is the extent to which our measure of GDV relates to other concepts used in sociology or in psychology, such as the fixed *versus* growth mindsets, or the Implicit Association Test measure. These concepts could help understanding why insisting more on positive behavior and effort aspects induces better performance compared to a more competence-oriented vocabulary.

Figure 3.1 – Share of Male and Female Students by Quartile of Percentile Rank at the DNB Math Exam



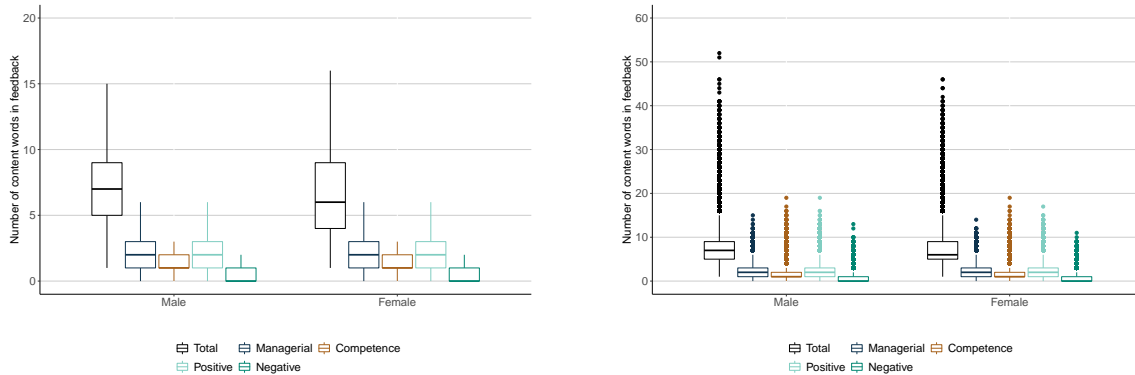
Notes: This graph shows the share of Grade 12 science major female and male students at each quartile of the math DNB percentile rank distribution, based on administrative data from the French Ministry of higher education.

Figure 3.2 – Enrollment in Higher Education by Gender after Grade 12 Science Track



Notes: This graph breaks down Grade 12 science major female and male students' matriculation choices into the main types of higher education programs, based on administrative data from the French Ministry of Education.

Figure 3.3 – Math Feedback - Distribution of Word Counts

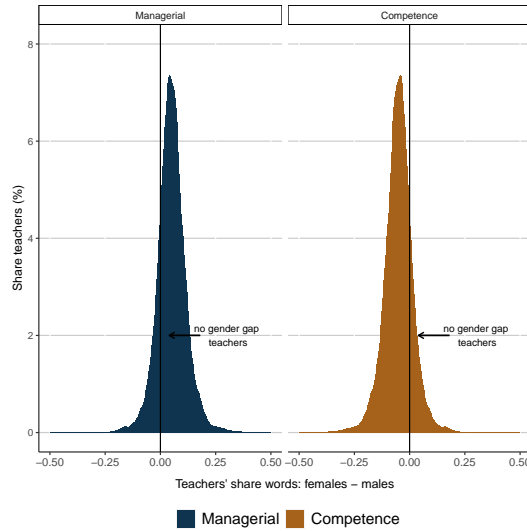


(a) Summary statistics

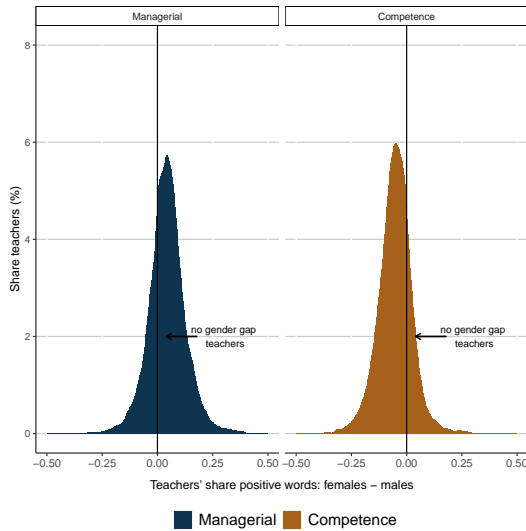
(b) Summary statistics and outliers values

Notes: This graph displays basic summary statistics on Grade 12 science major female and male students' distributions of feedback length in math, based on administrative data from the French Ministry of higher education. Each box displays the first and third quartile values as well as the median values. The segments cover the feedback length values that range between the first and third quartile values $\pm 1.5 \times \text{IQR}$, where IQR denotes the interquartile range. Values outside of this range can be considered as outlier values, and Panel (b) displays each of these outlier values separately with a dot.

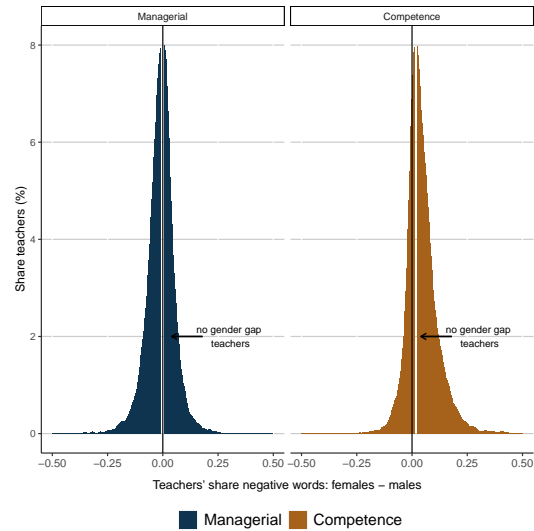
Figure 3.4 – Distribution of Teachers' Gender Gaps in Feedback Type and Positiveness - Females *minus* Males



(a) Teacher gender gap

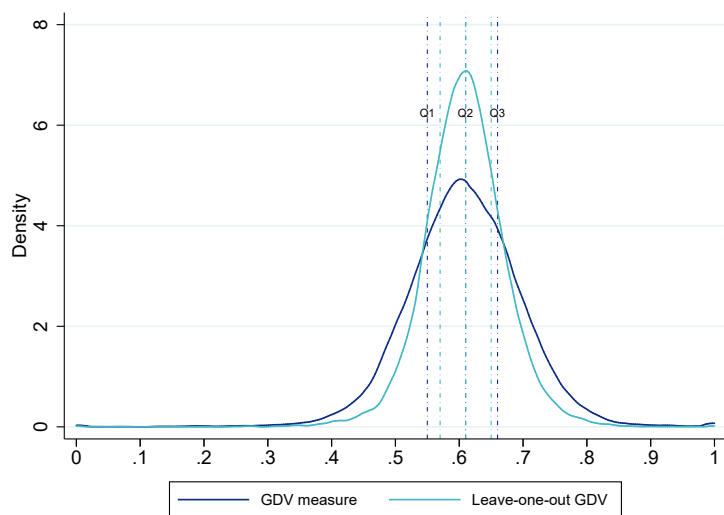
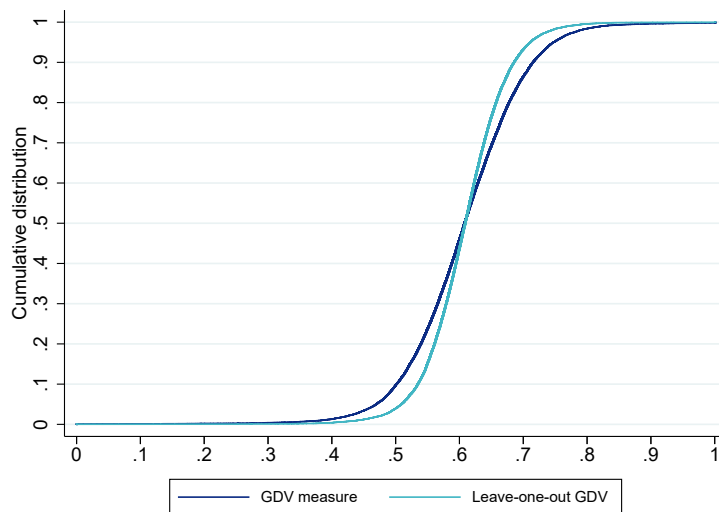


(b) Teacher gender gap in positive words



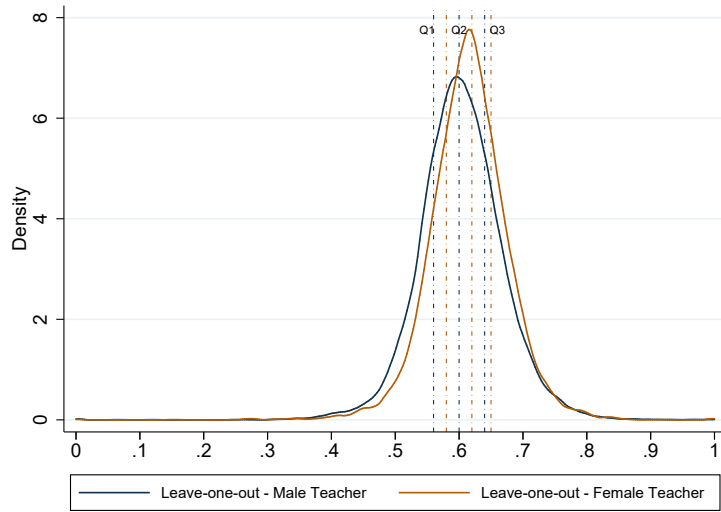
(c) Teacher gender gap in negative words

Notes: This figure shows the distributions of Grade 12 math teachers' gender gaps in feedback type and positiveness, based on administrative data from the French Ministry of higher education. Panel (a) displays the distributions of teachers' gender gaps in the share of Managerial and Competence words in the feedback given to their students. Panels (b) and (c) further detail the gender gaps distributions in the shares of positive and negative words, for each feedback type separately.

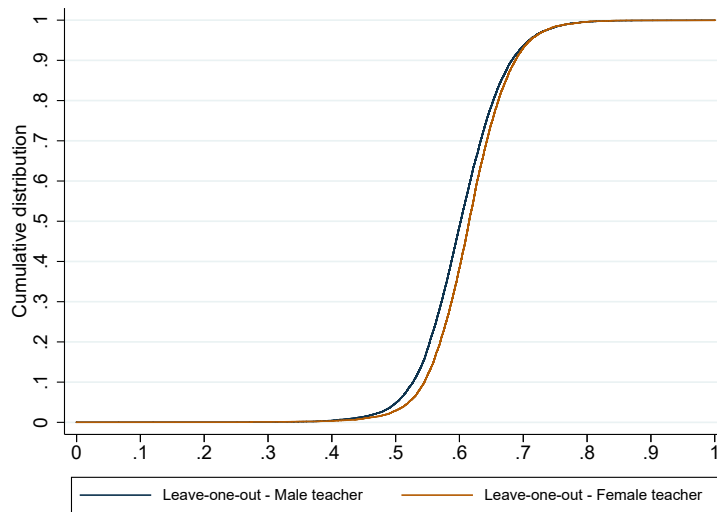
Figure 3.5 – Distribution of Math Teachers GDV and Leave-one-out GDV**(a)** Density**(b)** Cumulative distribution

Notes: This figure shows the distributions of math teachers' GDV and leave-one-out GDV measures, based on administrative data from the French Ministry of higher education. The sample consists of Grade 12 math teachers teaching in high school \times elective \times year cells containing more than one math teacher.

Figure 3.6 – Distribution of Math Teachers' Leave-one-out GDV – By Teacher Gender

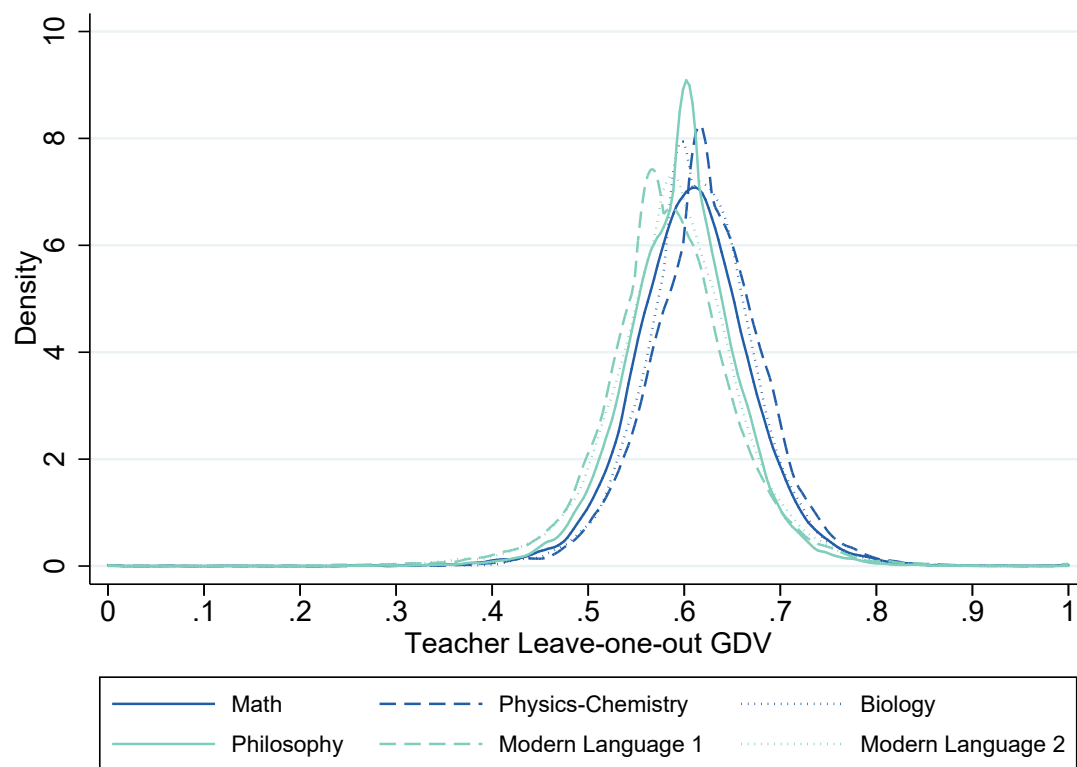


(a) Density



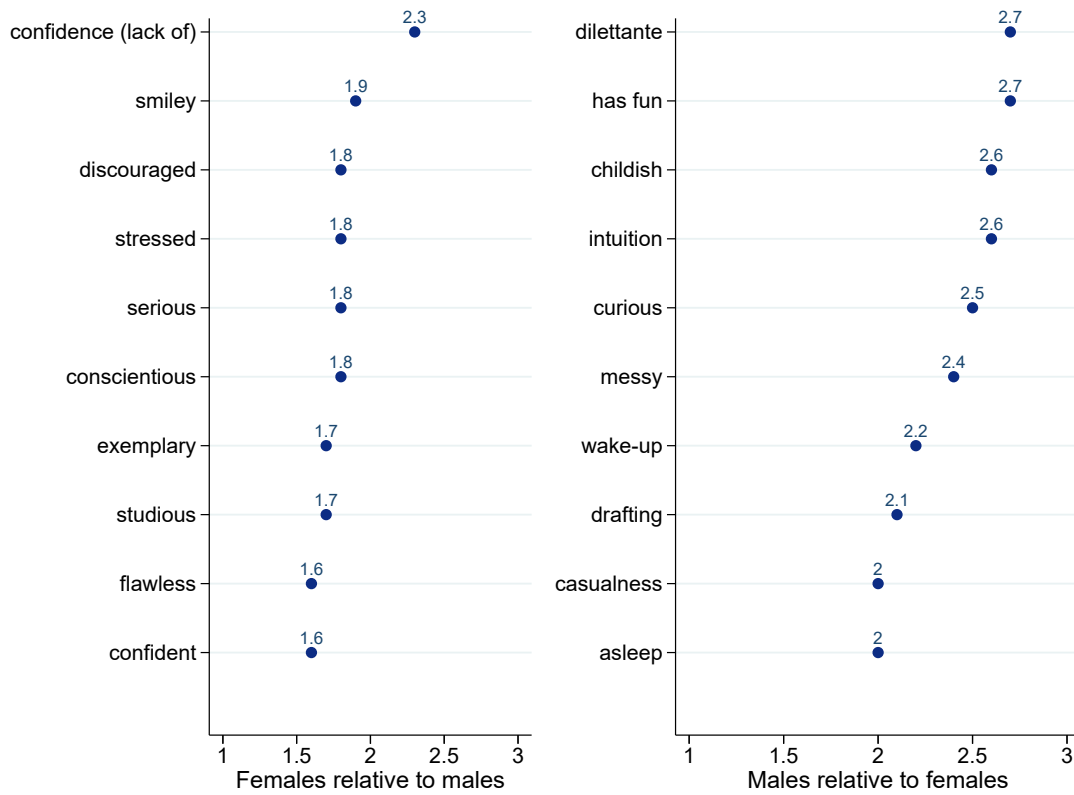
(b) Cumulative distribution

Notes: This figure shows the distributions of female and male math teachers' leave-one-out GDV measure, based on administrative data from the French Ministry of Education. The sample consists of Grade 12 math teachers teaching in high school \times elective \times year cells containing more than one teacher.

Figure 3.7 – Distribution of Teachers' Leave-one-out GDV – By Core Subjects

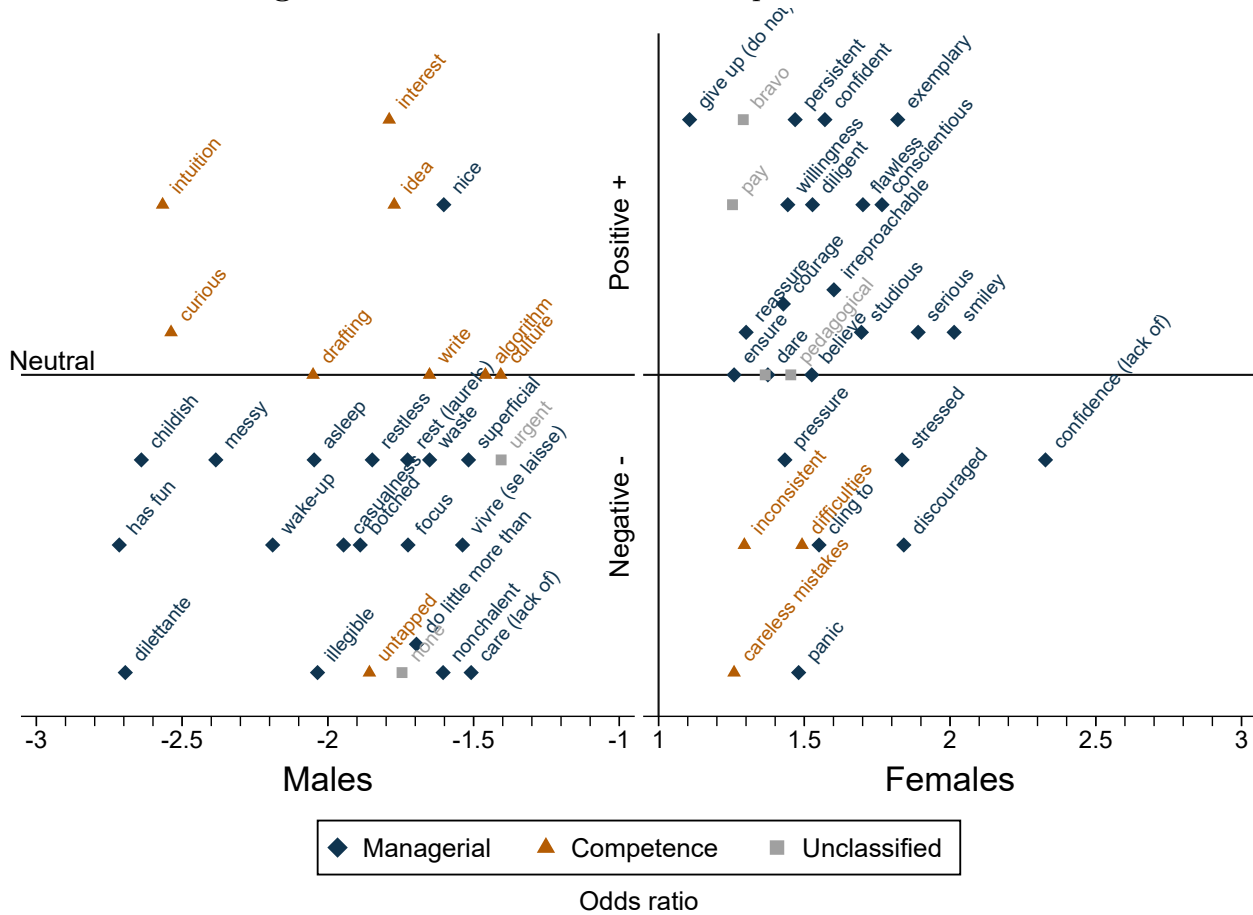
Notes: This figure shows the distributions of the math, physics, biology, philosophy and foreign language teachers' leave-one-out GDV measure, based on administrative data from the French Ministry of Education. The sample consists of Grade 12 teachers teaching in high school \times elective \times year cells containing more than one math teacher.

Figure 3.8 – Odds Ratios of the Top 10 Gender Predictors



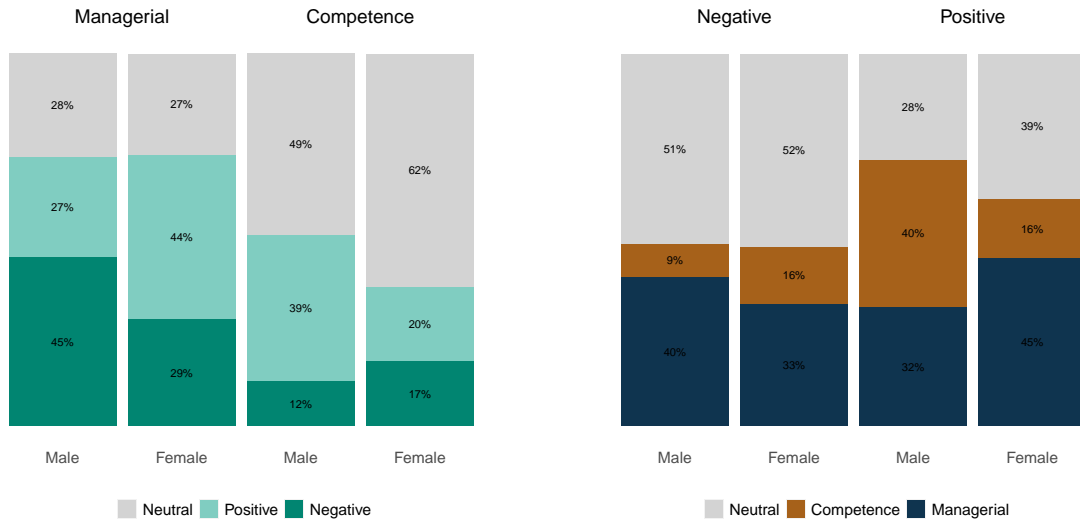
Notes: This figure shows the odds ratios obtained for the top 10 female and male predictors of the model described by Equation 3.1 estimated using the vocabulary appearing in math teachers' feedback. The estimation is realised on the universe of French Grade 12 science major students.

Figure 3.9 – Classification of the Top 30 Gender Predictors



Notes: This figure classifies the top 30 female and male predictors of the model described by Equation 3.1 estimated using the vocabulary appearing in math teachers' feedback into positive vs. negative and managerial vs. competence categories. Ambiguous words (i.e. the ones used in both positive and negative contexts) or words that do not fit in any of the categories are respectively labelled neutral or unclassified. The x-axis gives the odds-ratio of each predictor. The estimation is realised on the universe of French Grade 12 science major students.

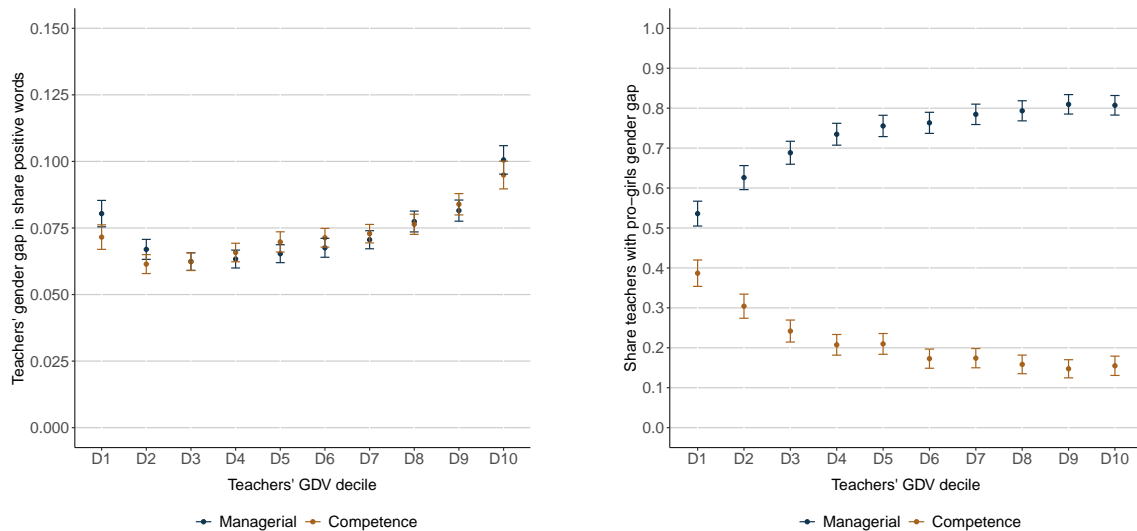
Figure 3.10 – Gender Predictors' Type and Positiveness



(a) Positiveness conditional on feedback type **(b)** Feedback type conditional on positiveness

Notes: This figure shows the proportions of managerial and competence-related gender predictors conditional on positiveness (Panel a), and that of positive and negative gender predictors conditional on feedback type (Panel b).

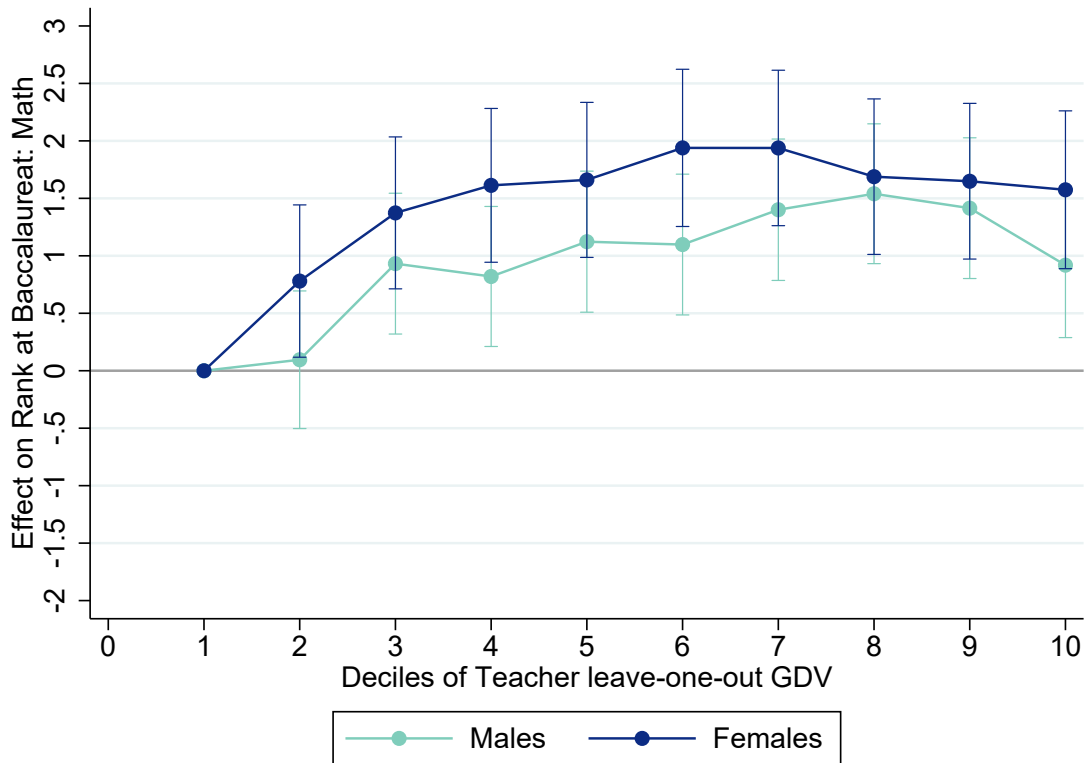
Figure 3.11 – Teachers' Gender Gap in the Share of Positive Words in Favor of Females by Deciles of GDV - In Absolute Value and Percentage



(a) Teacher's gender gap (absolute value) **(b)** Share of teachers with gender gap in favor of females

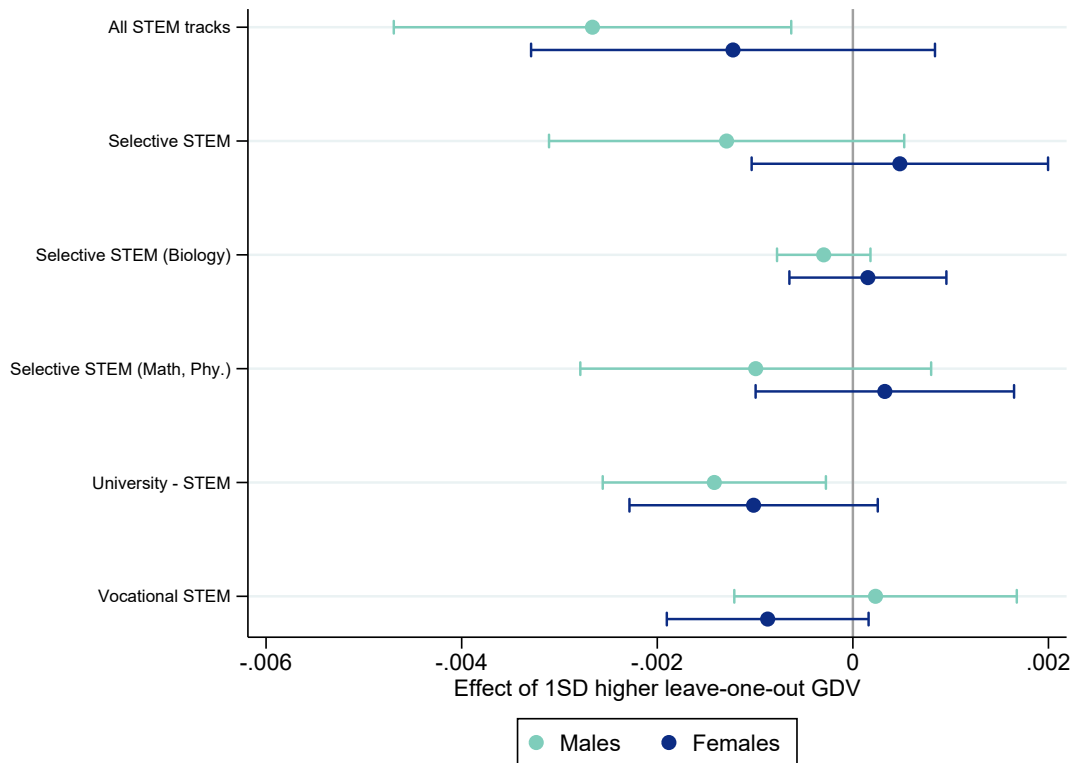
Notes: For each GDV decile, Panel (a) displays the average absolute value of Grade 12 teachers' gender gaps in the share of positive words appearing in their feedbacks, separately for competence vs. managerial related words. The GDV deciles are computed based on the teacher-level average of GDV measures. Panel (b) displays the share of teachers for whom the gender gap is in favour of female students, by GDV decile. The average values per decile are computed on the universe of math Grade 12 teachers for whom at least one GDV measure was estimated.

Figure 3.12 – Effect of Math Teacher GDV on Math Performance at *Baccalauréat* - By GDV Deciles



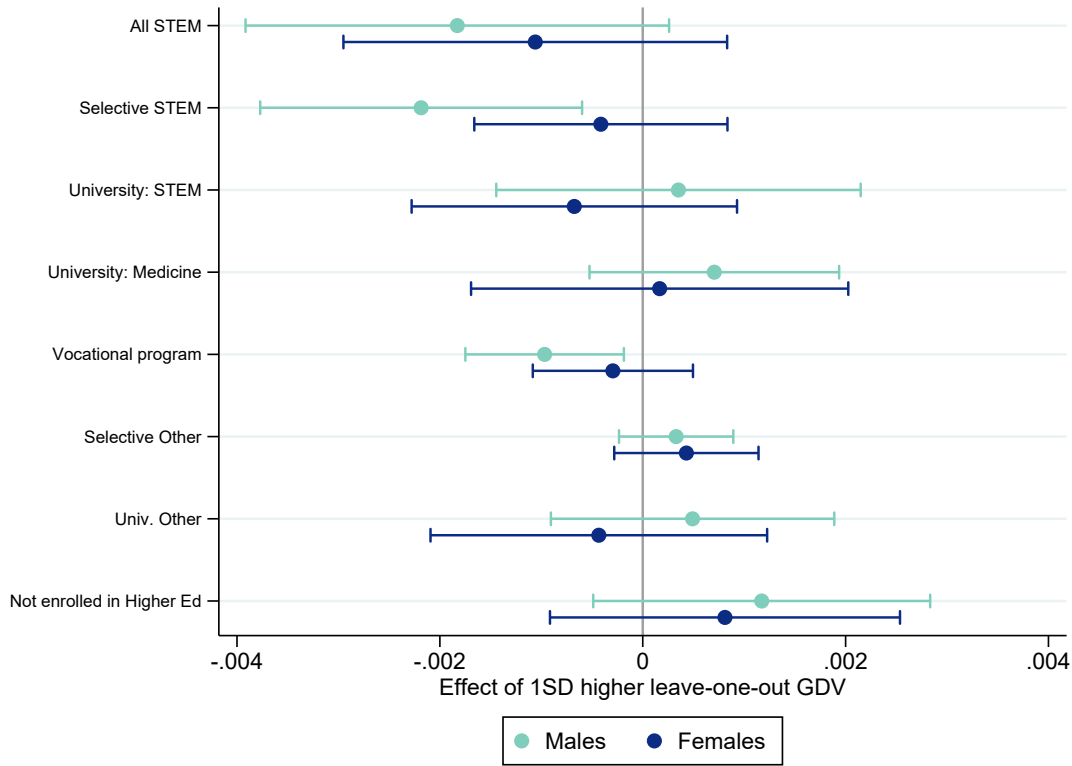
Notes: The results are calculated with administrative data from the French Ministry of higher education on French Grade 12 science major students. The figure reports the results of the regression of students' percentile rank at the math *baccalauréat* exam on a set of teacher leave-one-out GDV decile dummies, controlling for high school, year and elective fixed effects. Coefficients are expressed in deviation from the first decile's value, and are reported with their 95% confidence intervals. The coefficients are estimated on Grade 12 science major students for whom the high school \times elective \times year cell contains more than one math teacher.

Figure 3.13 – Effect of Math Teacher GDV on STEM Programs First Choice



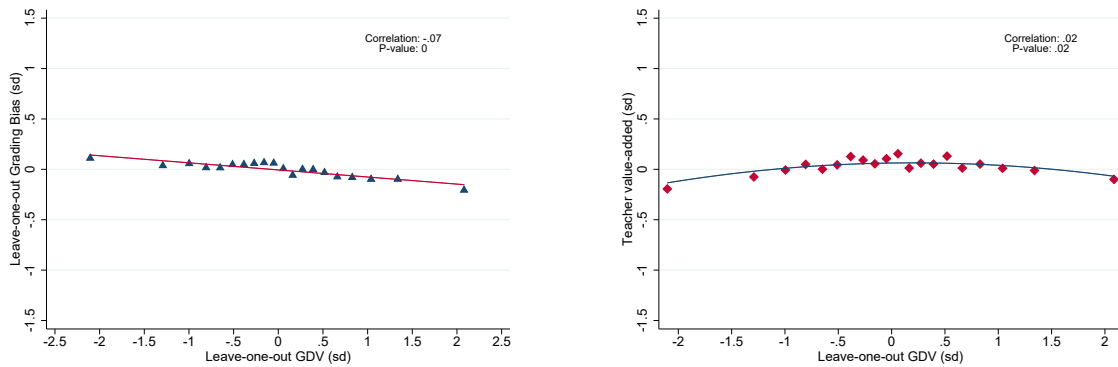
Notes: The results are calculated with administrative data from the French Ministry of higher education on French Grade 12 science major students. The figure shows the effect on female and male students' probability to top-rank STEM programs of being assigned a teacher with a one standard deviation higher leave-one-out GDV, along with its 95% confidence interval. The coefficients are estimated on students for whom the high school \times elective \times year cell contains more than one math teacher.

Figure 3.14 – Effect of Math Teacher GDV on Matriculation in the Following Year



Notes: The results are calculated with administrative data from the French Ministry of Education on French Grade 12 science major students. The figure shows the effect on female and male students’ probability of matriculation in the following year of being assigned a teacher with a one standard deviation higher leave-one-out GDV. The coefficients are estimated on students for whom the high school \times elective \times year cell contains more than one math teacher.

Figure 3.15 – Correlation Between Teacher GDV, Grading Bias and Teacher Quality



(a) Teacher GDV and Teacher Grading Bias

(b) Teacher GDV and Teacher Value-Added

Notes: The results are calculated with administrative data from the French Ministry of higher education on French Grade 12 science major students. The figure shows the binned average of the teachers’ leave-one-out grading bias (resp. value-added) standardised measures on the standardised teacher leave-one-out GDV. The line represents the linear fit in Panel (a) and the quadratic fit in Panel (b). The correlation coefficients are obtained from the regression of the grading bias (resp. value added) on teacher GDV. The sample consists of all Grade 12 math teachers for whom a leave-one out GDV measure, a leave-one-out grading bias measure and a value-added measure could be estimated.

Table 3.1 – Number of Grade 12 Students and Sample Restrictions

	2012	2013	2014	2015	2016	2017
Total nb. of G12 students	174,996	179,625	183,693	190,980	198,573	203,262
Nb. of obs with missing transcript	90,299	79,226	54,248	42,502	34,064	28,599
<i>% high school entirely missing:</i>	95.7	92.7	91.0	85.2	78.2	67.8
High school < 2 classes	3,634	4,716	6,443	6,836	7,516	7,680
Teachers < 2 classes	14,165	5,744	5,930	5,864	8,860	32,089
Obs. in the analytical sample	66,898	89,939	117,072	135,778	148,133	134,894
<i>(in %)</i>	(38.2)	(50.1)	(63.7)	(71.1)	(74.6)	(66.4)

Notes: This table reports the number of Grade 12 students appearing each year in the APB database. We show the number of observations removed for each sample restriction, and provide the number of observations used in the analytical sample in bold in the table. “High school entirely missing” refers to students enrolled in high schools that do not report grade transcripts automatically on the APB platform and that are therefore discarded from the sample.

Table 3.2 – Students’ Summary Statistics

	All	Males	Females
Demographics			
Female student (N= 691,093)	0.47		
Age (years) (N= 691,093)	18.09	18.12	18.06
Free lunch student (N= 691,059)	0.13	0.12	0.14
High SES (N= 691,093)	0.43	0.44	0.41
Medium-high SES (N= 691,093)	0.16	0.16	0.16
Medium-low SES (N= 691,093)	0.24	0.24	0.25
Low SES (N= 691,093)	0.17	0.16	0.18
Education			
Rank at DNB: math (N= 654,958)	50.28	52.18	48.13
Rank at DNB: French (N= 654,927)	50.33	44.69	56.73
Rank at <i>baccalauréat</i> : French (written) (N= 659,291)	50.00	45.01	55.62
Rank at <i>baccalauréat</i> : French (oral) (N= 659,254)	49.78	45.69	54.39
Maths elective (N= 622,903)	0.23	0.27	0.19
Physics-Chemistry elective (N= 622,903)	0.26	0.27	0.25
Earth & Life Science elective (N= 622,903)	0.37	0.26	0.50
Engineering & Info elective (N= 622,903)	0.13	0.20	0.06
Nb of classmates (N= 691,093)	30.83	30.68	31.01
Nb. of observations	691,093	368,922	322,171

Notes: This table shows descriptive statistics for the Grade 12 students in the analytical sample overall, and separately for males and females. The number of non-missing observations is reported in parentheses.

Table 3.3 – Teachers' Summary Statistics

	Mean	S.D
Share of head teacher at least once (N= 6,754)	0.53	0.50
Male math teacher (N= 6,721)	0.58	0.49
Number of teacher observations (N= 6,772)	3.70	1.64
Average number of classes per year (N= 6,772)	1.09	0.26
Average number of students per class (N= 6,772)	28.02	5.22
Average feedback length (N= 6,759)	7.40	2.51
Nb. of teachers	6,772	

Notes: This table shows descriptive statistics for the Grade 12 math teachers in the analytical sample. The average feedback length is computed as the average number of words in teachers' feedback, once common words (such as *the*, *she*, *a*, etc.) have been removed. The number of non-missing observations is reported in parentheses.

Table 3.4 – Balancing Test: Teachers’ Leave-One-Out GDV with Students’ Baseline Characteristics

	Dep. var: Teacher’s leave-one-out GDV (sd)		
	Coeff.	S.e	p-value
Female student	−0.0028	0.0028	0.3184
Age (years)	−0.0043*	0.0026	0.0994
Free lunch student	0.0041	0.0037	0.2719
Foreign student	0.0055	0.0087	0.5269
High SES	−0.3163	0.3123	0.3112
Medium-high SES	−0.3155	0.3124	0.3125
Medium-low SES	−0.3193	0.3122	0.3063
Low SES	−0.3174	0.3124	0.3096
Rank at DNB: math	−0.0001	0.0000	0.2922
Rank at DNB: French	0.0000	0.0001	0.5023
Rank at <i>baccalauréat</i> : French (written)	0.0001	0.0001	0.2298
Rank at <i>baccalauréat</i> : French (oral)	0.0001	0.0001	0.2421
High school, elective, year FE	Yes		
F-stat (p-value)	1.12	(0.334)	
Nb. of observations	573,595		

Notes: This table reports the estimation results of the teachers’ standardized leave-one-out GDV measure, defined at the class-level, regressed on the students’ socio-economic characteristics and baseline academic performance. The regression includes high school, year and elective course fixed-effects. Standard errors are clustered at the teacher level and are reported in the second column. The F-statistics test for the joint significance of regressors. ***: p-value < 0.01; **: p-value < 0.05, *: p-value < 0.1.

Table 3.5 – Pearson's Chi Square Tests of Class Random Assignment

	Nb. of nonmissing	Nb. of significant	Share sig. at	
	p-values	p-values at 5%	5%	1%
	(1)	(2)	(3)	(4)
Female	21,935	2,483	11.32	3.37
Age (years)	19,800	1,591	8.04	2.60
Free-lunch	19,170	973	5.08	1.29
Foreign student	8,094	328	4.05	1.36
High SES	22,151	1,489	6.72	1.44
Medium-high SES	20,847	929	4.46	0.78
Medium-low SES	21,928	1,156	5.27	0.93
Low SES	20,394	1,127	5.53	1.15
Rank at DNB: math	22,485	1,377	6.12	1.21
Rank at DNB: French	22,489	1,534	6.82	1.34
Rank at <i>baccalauréat</i> : French (written)	22,482	1,640	7.29	1.58
Rank at <i>baccalauréat</i> : French (oral)	22,475	1,613	7.18	1.41

Notes: This table reports the results of the Pearson Chi-square tests of independence performed on the 27,688 unique combinations of high schools, elective course and year. For each unique combination, we tabulate math teachers' identifiers with each baseline characteristic. Continuous variables such as age and percentile ranks are first discretized. Columns 3 and 4 report the share of p-values that are above the nominal levels of 5% and 1% respectively.

Table 3.6 – Impact of Teacher GDV on Academic Performance, Higher Education Choices and Enrollment

	All (1)	Males (2)	Females (3)
Academic performance			
Rank at <i>baccalauréat</i> : math	0.3351*** (0.0668)	0.3113*** (0.0747)	0.3735*** (0.0811)
Rank at <i>baccalauréat</i> : philo	-0.0235 (0.0646)	-0.0381 (0.0746)	-0.0093 (0.0786)
Type of STEM programs ranked first in the ROL			
All STEM tracks	-0.0019** (0.0008)	-0.0027** (0.0010)	-0.0012 (0.0011)
Selective STEM	-0.0005 (0.0007)	-0.0013 (0.0009)	0.0005 (0.0008)
<i>among which: biology</i>	-0.0001 (0.0002)	-0.0003 (0.0002)	0.0002 (0.0004)
<i>among which: math, physics</i>	-0.0004 (0.0007)	-0.0010 (0.0009)	0.0003 (0.0007)
University - STEM	-0.0012*** (0.0004)	-0.0014** (0.0006)	-0.0010 (0.0006)
Vocational STEM	-0.0003 (0.0005)	0.0002 (0.0007)	-0.0009* (0.0005)
Matriculation in the following year			
All STEM	-0.0014 (0.0008)	-0.0018* (0.0011)	-0.0011 (0.0010)
Selective STEM	-0.0013** (0.0006)	-0.0022*** (0.0008)	-0.0004 (0.0006)
University STEM	-0.0001 (0.0007)	0.0004 (0.0009)	-0.0007 (0.0008)
University Medicine	0.0004 (0.0006)	0.0007 (0.0006)	0.0002 (0.0009)
Vocational program	-0.0007** (0.0003)	-0.0010** (0.0004)	-0.0003 (0.0004)
Selective Other	0.0004* (0.0002)	0.0003 (0.0003)	0.0004 (0.0004)
University Other	0.0001 (0.0006)	0.0005 (0.0007)	-0.0004 (0.0008)
Not enrolled in Higher Ed	0.0010 (0.0007)	0.0012 (0.0008)	0.0008 (0.0009)
Nb. observations	649,105	345,201	303,904

Notes: Each row reports the coefficients of the standardized *leave-one-out* teacher GDV obtained from the estimation of Equation 3.6 for the different outcomes listed on the first column. It is estimated on the entire sample and separately for Grade 12 males and females. The regression includes high school, year and elective course fixed-effects. Standard errors are clustered at the teacher level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, *; p-value < 0.1.

Table 3.7 – Impact of Teacher GDV - Mechanisms

	Males		Females	
	Grading bias (1)	Value added (2)	Grading bias (3)	Value added (4)
Academic performance				
Rank at Baccalaureat: Math	0.3714*** (0.0770)	0.2486*** (0.0684)	0.3967*** (0.0851)	0.2825*** (0.0727)
<i>Coeff. on grading bias or VA</i>	0.5593*** (0.0788)	2.8396*** (0.1420)	-0.1437* (0.0861)	3.6506*** (0.1503)
Type of STEM programs ranked first in the ROL				
All STEM tracks	-0.0029*** (0.0011)	-0.0027** (0.0011)	-0.0015 (0.0011)	-0.0015 (0.0010)
<i>Coeff. on grading bias or VA</i>	-0.0010 (0.0011)	0.0014 (0.0018)	-0.0002 (0.0011)	-0.0013 (0.0016)
Selective STEM	-0.0015 (0.0010)	-0.0014 (0.0010)	0.0003 (0.0008)	0.0003 (0.0008)
<i>Coeff. on grading bias or VA</i>	-0.0004 (0.0010)	0.0036** (0.0017)	0.0001 (0.0008)	0.0004 (0.0013)
University - STEM	-0.0015*** (0.0006)	-0.0014** (0.0006)	-0.0010 (0.0006)	-0.0010 (0.0006)
<i>Coeff. on grading bias or VA</i>	-0.0003 (0.0006)	-0.0021** (0.0010)	-0.0008 (0.0007)	-0.0020* (0.0010)
Matriculation in the following year				
All STEM	-0.0022** (0.0011)	-0.0023** (0.0011)	-0.0014 (0.0010)	-0.0014 (0.0010)
<i>Coeff. on grading bias or VA</i>	-0.0010 (0.0011)	0.0051*** (0.0018)	-0.0003 (0.0010)	0.0010 (0.0015)
Selective STEM	-0.0025*** (0.0008)	-0.0025*** (0.0008)	-0.0004 (0.0007)	-0.0005 (0.0007)
<i>Coeff. on grading bias or VA</i>	0.0015* (0.0009)	0.0038*** (0.0015)	0.0002 (0.0007)	0.0003 (0.0011)
University: STEM	0.0002 (0.0010)	0.0002 (0.0010)	-0.0010 (0.0008)	-0.0010 (0.0008)
<i>Coeff. on grading bias or VA</i>	-0.0023** (0.0010)	0.0015 (0.0015)	-0.0005 (0.0009)	0.0007 (0.0013)
Nb. observations	340,729	342,380	300,020	301,409

Notes:

Table 3.8 – Impact of Teacher GDV-Females and GDV-Males - Mechanisms

	All		Males		Females	
	GDV Males	GDV Females	GDV Males	GDV Females	GDV Males	GDV Females
	(1)	(2)	(3)	(4)	(5)	(6)
Academic performance						
Rank at <i>baccalauréat</i> : math	0.3921*** (0.0854)	0.4905*** (0.0904)	0.4497*** (0.0953)	0.4828*** (0.0986)	0.3358*** (0.1040)	0.5232*** (0.1110)
Type of STEM programs ranked first in the ROL						
All STEM tracks	-0.0017 (0.0011)	-0.0022** (0.0011)	-0.0028** (0.0014)	-0.0029** (0.0014)	-0.0007 (0.0013)	-0.0014 (0.0013)
Selective STEM	-0.0010 (0.0009)	0.0004 (0.0009)	-0.0019 (0.0012)	-0.0005 (0.0013)	0.0002 (0.0010)	0.0016 (0.0010)
University STEM	-0.0003 (0.0006)	-0.0020*** (0.0006)	-0.0006 (0.0008)	-0.0025*** (0.0008)	-0.0002 (0.0008)	-0.0016* (0.0008)
Matriculation in the following year						
All STEM	-0.0017 (0.0011)	-0.0014 (0.0011)	-0.0024* (0.0014)	-0.0017 (0.0014)	-0.0011 (0.0012)	-0.0013 (0.0012)
Selective STEM	-0.0021*** (0.0008)	-0.0001 (0.0008)	-0.0035*** (0.0011)	-0.0008 (0.0011)	-0.0005 (0.0009)	0.0005 (0.0008)
University STEM	0.0004 (0.0009)	-0.0014 (0.0009)	0.0012 (0.0012)	-0.0010 (0.0012)	-0.0006 (0.0010)	-0.0019* (0.0010)
Nb. observations	642,584	642,433	341,168	341,111	301,416	301,322

Notes: Each row reports the coefficient of the standardized *leave-one-out* teacher GDV-males and GDV-females measures obtained from the estimation of Equation 3.6 for the different outcomes listed on the first column. It is estimated on all the sample and separately for Grade 12 males and females. The regression includes high school, year and elective course fixed-effects. Standard errors are clustered at the teacher level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05, ***; p-value < 0.1.

Appendix to
Teacher Gendered Feedback, Students' Math Performance
and Enrollment Outcomes

3.A Measuring Teachers' GDV: Details of the Estimation Procedure

This appendix provides the details on the practical implementation for the different steps of teachers' gender differentiated feedback (GDV) estimation procedure developed in Section 3.3.

3.A.1 Textual Data Preparation

The students academic records consist of a corpus of *documents*, where a *document* corresponds to the feedback that a teacher gave to a given student, in a given subject. Our aim is to convert all the documents into a data structure similar to the one displayed in Table 3.A1. In this example, all the words and grouping of two words that appear at least once in a document have been converted to a column.

Text cleaning. In order to reduce the dimensionality of our data and, consequently, the computational burden of our estimation, we follow the text cleaning steps suggested by Gentzkow et al. (2019). For each *document*, we remove all punctuation signs, but keep track of the position of full stops in order to identify the different sentences that composed the original text. We get rid of all first names (that are identified based on the Insee register of French first names), which would be very good predictors of student gender without reflecting any gender differentiation of the vocabulary used. We also remove *stop words*, which are very common words that bear little informational content, like “*le*” (“the”), “*donc*” (“thus”), “*déjà*” (“already”), etc...

All remaining words are *stemmed*, i.e., replaced by their roots: for instance, the words “*amateur*” and “*amatrice*” are replaced by their common root “*amat*”. This last step is crucial to our analysis, because it allows to get rid of all the grammatical markers of the students' gender, which often appear, in French, at the end of the words. We further reduce the dimensionality of our data by getting rid of all *stemmed* words that appear in less than 100 documents.

Tokenization. In order to convert the remaining words into a set of columns (also known as the document-term matrix), we “dummify” words and grouping of words. Each word that appears in the corpus becomes a column, that takes value one if the word appears in the document, and zero otherwise. In the text analysis literature, groups of words are commonly denoted *ngrams*, where n corresponds to the number of words in the considered group of words. In our analysis, we choose

to use as regressors *unigrams*, i.e., tokens composed of only one word.

Table 3.A1 – From Text to Data: An Illustration

Document	ensemble	alarmant	bon	travail	sérieux	ensemble alarmant	bon travail
<i>Ensemble alarmant, manque de sérieux.</i>	1	1	0	0	1	1	0
<i>Bon travail, beaucoup de sérieux.</i>	0	0	1	1	1	0	1

Token classification. Building on the considerable literature in psychology and sociology related to the analysis of gendered feedback (Dweck et al., 1978; Morgan, 2001), we classify the gender predictors into one of the four following categories: positive (resp. negative) competence-related aspects and positive (resp. negative) managerial aspects. We classify as competence-related any word that either relates to math and the school environment (e.g. reasoning, exam, geometry) or to an intellectual skill (e.g. talented, potential). Ambiguous words (i.e., the ones used in both positive and negative contexts) or words that do not fit in any of the categories are respectively labeled neutral or unclassified. The words classified as managerial are the adjectives used to describe the student’s behavior in class (e.g. shy, exemplary) as well as the actions undertaken or the efforts provided by the student (e.g. involved, revise, fall behind). The classification of the 100 best predictors of each gender can be found in Tables 3.A2 and 3.A3.^{A.1A.2}

3.A.2 Predicting Student Gender and Measuring Teacher GDV

In this second step, the tokens are used as predictors of students gender. We assume that the probability of being a female student conditional on the words used in the feedback has a logistic

^{A.1}Even though every token has been classified, we only show the top 100 predictors given that other predictors are not more frequently used for female or male students (their odds ratio is around 1) and cannot be classified in any of the four mentioned categories in the vast majority of cases.

^{A.2}We also attempted to build these categories in a data-driven manner using bi-term topic models tailored for short texts, but these models performed poorly on our data. Our data is indeed quite specific in that texts are very short, with an average number of tokens equal to 7, the overall vocabulary is quite limited ($\simeq 1,600$ words) with little variation in the topics used as they all relate to academic performance and behavior. We therefore faced the typical challenges inherent to such short texts: the generated topics gathered inconsistent words (*trivial topics*) and the different topics were highly similar with a lot of words in common (*repetitive topics*, see Wu et al. (2020) for a discussion on those issues.)

form:

$$P(Female_i = 1|W_i) = \frac{\exp(\alpha W_i)}{1 + \exp(\alpha W_i)} \quad \forall i \quad (\text{A.1})$$

and our objective is to find the set of α coefficients that maximize the penalized log-likelihood function, where λ is the regularization parameter:

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}}(-\ln(L(\alpha)) + \lambda \sum_{w=1}^{W_n} |\alpha_w|) \quad (\text{A.2})$$

The $\hat{\alpha}$ estimates are then used to predict students' gender. The teacher's GDV measure is computed based on those predictions, and is defined as the proportion of students for whom the model correctly predicts their gender, separately for each teacher. In practice, we estimate a logistic-Lasso to determine the α coefficients. We detail below the practical implementation of the estimation.

Step 1: Undersampling. Before any estimation is done, we deal with the issue of gender imbalance using undersampling techniques. Because gender is correlated with math performance, the model is likely to perform better on classes with larger gender imbalances in terms of math performance. For that reason, for each class, we sample as many males and females from each quartile of prior math performance. We define quartiles of prior math performance as follows: within each Grade 12 class, we rank students according to the math grade obtained at the DNB exam and create quartiles. Then for each class \times quartile we select n_{cq} males and n_{cq} females where $n_{cq} = \min(n_{cq}^{females}; n_{cq}^{males})$.^{A.3}

Step 2: Random selection of tokens. As shown in Table 3.3, the number of tokens used in feedback varies by teacher. As feedback length could influence the quality of the prediction, we randomly sample tokens for lengthy feedback, defined as the ones with an above-median length. For such feedback we randomly select six tokens, which is the median number.

Step 3: Training and hold-out samples. To avoid overfitting concerns, we fit the model described by Equation (A.2) on a training sample (30 percent of the undersampled data) and predict gender on a hold-out sample (70 percent). To preserve the balanced structure of the

^{A.3}We use the French grade obtained at the DNB exam instead of the math grade when we compute the teacher GDV for humanities related subjects, i.e., for philosophy and modern languages.

undersampled data, the partition of the data into a training and a hold-out sample is stratified, i.e., we include 30 percent (70 percent) of n_{cq} males and females in the training (hold-out) sample.

Step 4: Training the model. The training sample is used to fit the model and get the estimated $\hat{\alpha}$ coefficients. We first tune the regularization parameter λ by running a logistic Lasso with a 10-fold cross validation. We pick the λ value that lies within one standard deviation of the minimal error (Hastie et al., 2009) and estimate the logistic-lasso to obtain the $\hat{\alpha}$.

Step 5: Predict students' gender. The fitted model is applied to the hold-out sample to predict each student's gender. The model classifies a student as a girl ($\widehat{Sex}_i = 1$) if the predicted probability is greater than 0.5, and as boy otherwise ($\widehat{Sex}_i = 0$).

Step 6: Compute the teacher GDV measure. Finally for each class c of teacher j , we compute the GDV measure as the average proportion of correctly classified students:

$$GDV_{jc} = \frac{1}{N_{jc}} \sum_{i=1}^{N_{jc}} \mathbb{1}\{Sex_i = \widehat{Sex}_i\} \times 100 \quad \forall j, c \quad (\text{A.3})$$

where N_{jc} is the number of students in the balanced subsample of teacher j 's students from class c :

$$N_{jc} = \sum_{q=1}^4 2 \times n_{cq}$$

The teacher GDV measure defined by Equation A.3 could capture some unobserved-class specific gender differences. To rule out this concern, we also compute the *leave-one-out* teacher GDV as the average GDV over all the other classes taught except the current one:

$$GDV_{j \setminus c} = \frac{1}{N_j - 1} \sum_{c' \neq c} GDV_{jc'} \quad \forall j, c \quad (\text{A.4})$$

The two GDV measures are inherently noisy as they are computed on a limited number of observations. To stabilize those two measures and in order for our results not to depend on a single data split defined at Step 2, we repeat Step 1 to Step 5 100 times and use the GDV measures averaged over those 100 iterations.

Table 3.A2 – Top 100 Predictors' Classification - Female

	Positive	Negative	Neutral
Competence-related	3 tokens: autonomous, master, quality	5 tokens: careless mistakes, difficulties, inconsistent, mistake, misunderstandings	10 tokens: appropriate, calculus, classical, guidelines, method, question, read, support, usual
Managerial	29 tokens: cling to, confident, conscientious, courage, deserve, determined, diligent, discrete, efficient, encourage, exemplary, flawless, give up (do not), impeccable, irreproachable, keep doing, pay, persevere, persistent, reassure, reward, serious, seriously, smiley, steady, studious, tenacious, voluntary, willingness	15 tokens: cling to, concern, confidence (lack of), discouraged, doubt, give up, hesitate, panic, pressure, shy, stressed, suffer, unassuming, worry	6 tokens: believe, dare, ensure, intervene, pursue
Unclassified	6 tokens: bravo, congratulations, fruit (of work), laudable, pay, positive	5 tokens: decrease, drop, fragile, mishap, too low	22 tokens: a lot, allow, also, benchmark, big, circumstances, complete, context, contribute, despite, health, help, illustrate, know, other, pedagogical, point, pupil, recover, regular

Notes: This table reports the classification of the 100 best female feedback predictors into one of the four following categories: positive (resp. negative) competence-related aspects and positive (resp. negative) managerial aspects. We classify as competence-related any word that either relates to math and the school environment (e.g. reasoning, exam, geometry) or to an intellectual skill (e.g. talented, potential). The words classified as managerial are the adjectives used to describe the student's behavior in class (e.g. shy, exemplary) as well as the actions undertaken or the efforts provided by the student (e.g. involved, revise, fall behind). Ambiguous words (i.e. the ones used in both positive and negative contexts) or words that do not fit in any of the categories are respectively labelled neutral or unclassified.

Table 3.A3 – Top 100 Predictors’ Classification - Male

	Positive	Negative	Neutral
Competence-related	15 tokens: ambition, aptitude, capability, capacities, curious, idea, interest, intuition, passion, potential, relevant, rigorous, rigourous, scientific, sharpness	3 tokens: mix up, slow, untapped	13 tokens: algorithm, argument, computing, culture, drafting, expression (oral/written), guidelines, homework, passage, word, write, writing, written
Managerial	5 tokens: consciousness, detailed, nice, reaction, rebound	28 tokens: asleep, botched, care (lack of), casualness, childish, dilettante, disorganized, do little more than, focus, has fun, illegible, immature, inexistant, messy, minimal, nonchalant, rest (laurels), restless, scattered, shake up, skim through, superficial, troublesome, unacceptable, vivre (se laisse), wake-up, waste	5 tokens: behave, exploit, intervene, justify, work
Unclassified	3 tokens: best, easy, sufficient	8 tokens: excessive, insufficient, minimum, none, perfectible, shame, sufficient, urgent	21 tokens: advice, could, decide, double, expected, handed in, imposed, invite, mature, measure, obvious, outside, personal, put, radical, time, took, want

Notes: This table reports the classification of the 100 best male feedback predictors into one of the four following categories: positive (resp. negative) competence-related aspects and positive (resp. negative) managerial aspects. We classify as competence-related any word that either relates to math and the school environment (e.g. reasoning, exam, geometry) or to an intellectual skill (e.g. talented, potential). The words classified as managerial are the adjectives used to describe the student’s behavior in class (e.g. shy, exemplary) as well as the actions undertaken or the efforts provided by the student (e.g. involved, revise, fall behind). Ambiguous words (i.e. the ones used in both positive and negative contexts) or words that do not fit in any of the categories are respectively labelled neutral or unclassified.

3.B Teacher Gender Composition - All Core Subjects

Table 3.B1 – Share of Male Teachers by Core Subjects

Subject	Share	N	% non-miss
Math	0.58	7,124	0.93
Physics-Chemistry	0.57	7,760	0.93
Biology	0.37	6,694	0.92
Philosophy	0.62	7,420	0.95
Modern language 1	0.20	17,589	0.88
Modern language 2	0.19	22,613	0.83

Notes: This table reports the share of male teachers for each core subject taught to Grade 12 students in our sample. The second column gives the number of teachers and the third column gives the proportion of teachers for whom the sex was not missing.

3.C Assessing the Randomness of Missing Grade Transcripts

Table 3.C1 – Balancing Test: High Schools With All Missing Grade Transcripts

	Dep. var: Grade transcripts all missing in high school		
	Coeff.	S.e	p-value
Female student	−0.1167***	0.0362	0.0013
Age (years)	0.1875***	0.0141	0.0000
Free lunch student	−0.3854***	0.0485	0.0000
Foreign student	−0.0002	0.0871	0.9981
High SES	0.0110	0.0421	0.7938
Medium-high SES	−0.3664***	0.0691	0.0000
Medium-low SES	−0.1253**	0.0535	0.0191
Rank at DNB maths	0.0026	0.0021	0.2309
Rank at DNB math (females)	0.0011	0.0010	0.2599
Rank at DNB math (males)	−0.0039***	0.0013	0.0023
Nb. of observations	12,864		

Notes: This table reports the estimation results of dummies indicating whether the high school is systematically not reporting grade transcripts, regressed on the high school students' average characteristics. Standard errors are clustered at the high school level and are reported in the second column. ***: p-value < 0.01; **: p-value < 0.05, ***; p-value < 0.1.

3.D Robustness Checks and Additional Results

3.D.1 Robustness Checks

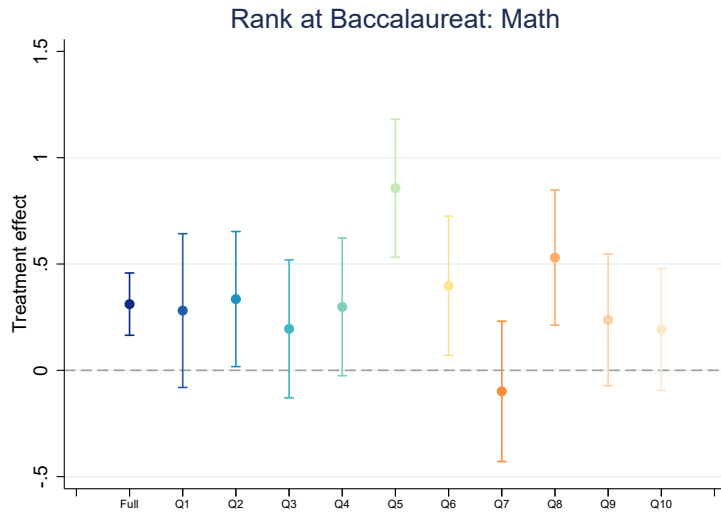
Table 3.D1 – Impact of Teacher GDV - Robustness Checks

	Males			Females		
	<i>Baseline</i> <i>X</i> (1)	<i>% of</i> <i>females</i> (2)	<i>Average</i> <i>GDV</i> (3)	<i>Baseline</i> <i>X</i> (4)	<i>% of</i> <i>females</i> (5)	<i>Average</i> <i>GDV</i> (6)
Academic performance						
Rank at <i>baccalauréat</i> : math	0.3541*** (0.0690)	0.3979*** (0.0744)	0.3985*** (0.0743)	0.4978*** (0.0730)	0.5027*** (0.0810)	0.5047*** (0.0811)
Rank at <i>baccalauréat</i> : philo	-0.0797 (0.0689)	-0.0233 (0.0745)	-0.0263 (0.0745)	-0.0691 (0.0694)	-0.0432 (0.0793)	-0.0455 (0.0793)
Type of STEM programs ranked first in the ROL						
All STEM tracks	-0.0025** (0.0011)	-0.0025** (0.0010)	-0.0023** (0.0010)	-0.0015 (0.0011)	-0.0013 (0.0011)	-0.0013 (0.0011)
Selective STEM	-0.0016* (0.0009)	-0.0013 (0.0009)	-0.0013 (0.0009)	0.0005 (0.0008)	0.0004 (0.0008)	0.0004 (0.0008)
<i>among which: biology</i>	-0.0004* (0.0003)	-0.0004 (0.0002)	-0.0004 (0.0002)	0.0001 (0.0004)	0.0000 (0.0004)	0.0000 (0.0004)
<i>among which: math, physics</i>	-0.0012 (0.0009)	-0.0010 (0.0009)	-0.0009 (0.0009)	0.0004 (0.0007)	0.0004 (0.0007)	0.0004 (0.0007)
University STEM	-0.0010* (0.0006)	-0.0012* (0.0006)	-0.0012* (0.0006)	-0.0014** (0.0007)	-0.0011 (0.0007)	-0.0011 (0.0007)
Vocational STEM	0.0003 (0.0008)	0.0002 (0.0007)	0.0002 (0.0007)	-0.0007 (0.0005)	-0.0008 (0.0005)	-0.0008 (0.0005)
Matriculation in the following year						
All STEM tracks	-0.0017 (0.0011)	-0.0016 (0.0010)	-0.0015 (0.0011)	-0.0012 (0.0010)	-0.0012 (0.0010)	-0.0012 (0.0010)
Selective STEM	-0.0025*** (0.0008)	-0.0023*** (0.0008)	-0.0022*** (0.0008)	-0.0002 (0.0006)	-0.0002 (0.0006)	-0.0002 (0.0006)
University STEM	0.0008 (0.0010)	0.0006 (0.0009)	0.0007 (0.0009)	-0.0010 (0.0009)	-0.0010 (0.0008)	-0.0010 (0.0008)
University Medicine	0.0007 (0.0007)	0.0007 (0.0006)	0.0007 (0.0006)	0.0002 (0.0010)	0.0002 (0.0010)	0.0002 (0.0010)
Vocational program	-0.0009** (0.0004)	-0.0009** (0.0004)	-0.0009** (0.0004)	-0.0003 (0.0004)	-0.0004 (0.0004)	-0.0004 (0.0004)
Selective Other	0.0004 (0.0003)	0.0004 (0.0003)	0.0004 (0.0003)	0.0004 (0.0004)	0.0005 (0.0004)	0.0005 (0.0004)
University Other	0.0005 (0.0008)	0.0002 (0.0007)	0.0002 (0.0007)	-0.0003 (0.0009)	-0.0004 (0.0008)	-0.0004 (0.0008)
Not enrolled in Higher Ed	0.0008 (0.0008)	0.0008 (0.0009)	0.0008 (0.0009)	0.0009 (0.0008)	0.0009 (0.0009)	0.0010 (0.0009)
Nb. observations	314,389	344,232	344,227	282,260	303,206	303,204

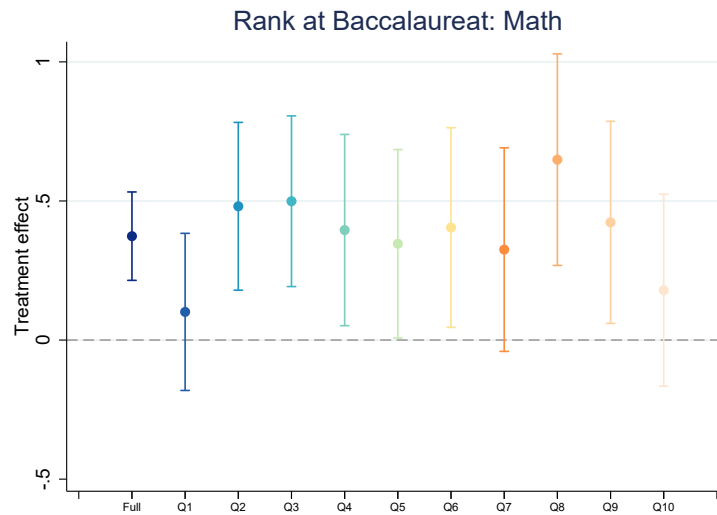
Each row reports the coefficients of the standardized *leave-one-out* teacher GDV obtained from the estimation of Equation 3.6 for the different outcomes listed on the first column. It is estimated on all the sample and separately for Grade 12 males and females. The regression includes high school, year and elective course fixed-effects. Columns 1 and 4 further control for the set of students' baseline characteristics listed in Table 3.2, columns 2 and 5 control for the average proportion of females in the classroom, and columns 3 and 6 control for the average leave-one-out GDV measured in other subjects for students from the same class. Standard errors are clustered at the teacher level and are reported in parentheses. ***: p-value < 0.01; **: p-value < 0.05; *: p-value < 0.1.

3.D.2 Additional Results: Heterogeneity by Initial Math Performance

Figure 3.D1 – Impact of Teacher GDV on Math Performance at *Baccalauréat* - By Deciles of Initial Math Performance



(a) Males

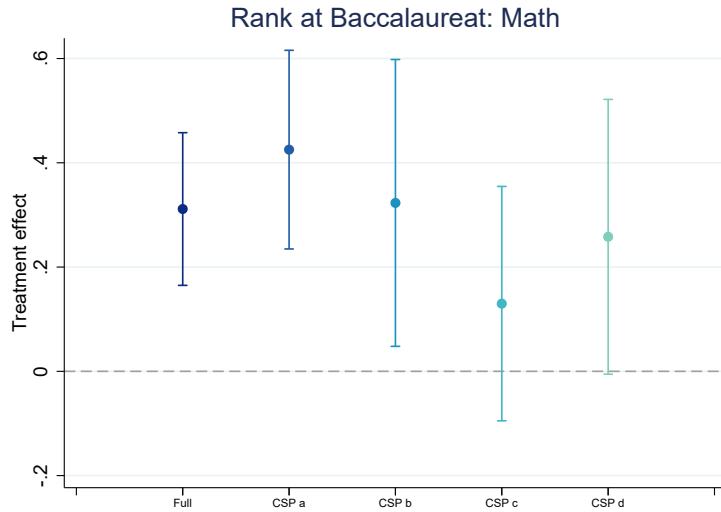


(b) Females

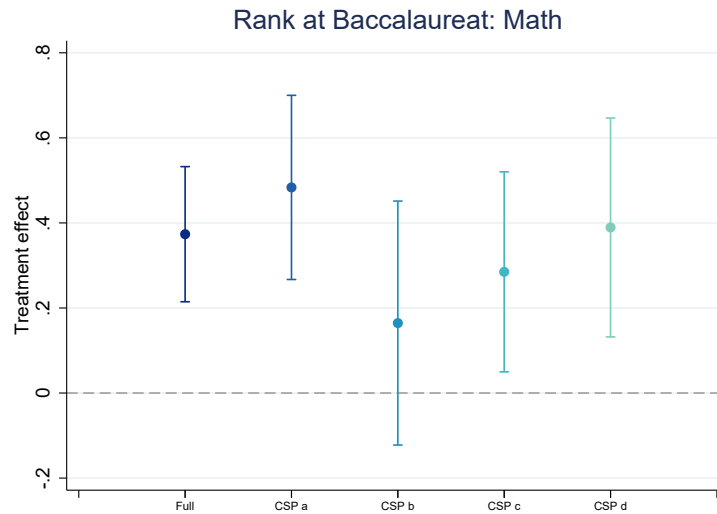
Notes: The figure reports the effect of a one standard deviation increase in teacher leave-one-out GDV on students' rank at *baccalauréat* in math separately by gender and by initial performance in math. Initial math performance is measured as deciles of percentile rank in math obtained at the DNB nation exam in Grade 9. The solid dots show the estimated coefficients, with 95 percent confidence intervals denoted by vertical capped bars. The coefficients are estimated on students for whom the high school \times elective \times year cell contains more than one math teacher.

3.D.3 Additional Results: Heterogeneity by Social Background

Figure 3.D2 – Impact of Teacher GDV on Math Performance at *Baccalauréat* - By Social Background



(a) Males



(b) Females

Notes: The figure reports the effect of a one standard deviation increase in teacher leave-one-out GDV on students' rank at *baccalauréat* in math separately by gender and by socioeconomic background. The solid dots show the estimated coefficients, with 95 percent confidence intervals denoted by vertical capped bars. The coefficients are estimated on students for whom the high school \times elective \times year cell contains more than one math teacher.

3.E Mechanisms: Estimation Details and Complementary Results

3.E.1 Estimating the Teacher Grading Bias

We follow Lavy and Sand (2018a) and Terrier (2020) and compute the teacher grading bias as the difference between the class gender gaps in the non-blind (NB) and blind scores (B). We use the (standardized) math grade obtained at the continuous assessment as the non-blind score, and the (standardized) math grade obtained the *baccalauréat* exam as the blind score. The grading bias (GB) for class c taught by teacher j in year t is therefore defined as follows:

$$GB_{cjt} = \left(NB_{cjt}^{males} - NB_{cjt}^{females} \right) - \left(B_{cjt}^{males} - B_{cjt}^{females} \right)$$

The grading bias assigned to class c is actually the average bias observed in any other classes taught by the same teacher except class c itself, i.e., it is the leave-one-out grading bias. A negative (positive) grading bias is indicative of a bias in favor of females (males).

The table below reports the average standardized non-blind and blind scores separately for Grade 12 males and females. We see that on average, females score above the mean class grade at the continuous assessment, but below when we consider the math *baccalauréat* grade. The reverse holds for males. The teacher grading bias is calculated as the difference between columns 3 and 6, and is negative, thus revealing a grading bias favoring females, both from male and female teachers.

3.E.2 Estimating the Teacher Value-Added

Teachers' value-added are estimated using the three steps described in the Chetty et al. (2014) paper. The steps are implemented using the `vam` package developed by Stepner (2013). We detail those three steps below.

Table 3.E1 – Teachers’ Average Math Grading Bias for Grade 12 Students

	Males			Females			Teacher
	G12 maths (1)	Bac maths (2)	Diff. (3)	G12 maths (4)	Bac maths (5)	Diff. (6)	bias
All teachers	−0.017	0.043	−0.060	0.020	−0.048	0.068	−0.129
Female teachers	−0.029	0.028	−0.057	0.033	−0.031	0.064	−0.121
Male teachers	−0.009	0.054	−0.063	0.010	−0.062	0.072	−0.135
N	364,611	343,945		319,499	306,551		

Notes: This table reports the average standardized math grades obtained at the Grade 12 continuous assessment (columns 1 and 4) and that obtained at the math *baccalauréat* exam (columns 2 and 4) separately for males and females. Columns 3 and 6 report the average difference between both grades. The teacher grading bias reported in the last column of the table reports the average grading bias computed at the teacher level, obtained as the difference between columns 3 and 4. A negative grading bias is indicative of bias in favour of girls.

Step 1: Residualizing students test scores. We first regress students’ test scores in year t , measured by the percentile rank obtained at the math *baccalauréat*, on a set of students’ baseline covariates, controls for students’ prior performance, previous year’s class characteristics, and teachers fixed effects.

- *Students’ baseline characteristics:* gender; free-lunch status; four dummies for students’ SES background (low SES, medium-low SES, medium-high SES, high SES); a dummy equal to one if the student is a foreigner.
- *Students’ prior performance:* It includes the math grade obtained during the Grade 11 continuous assessment, standardized by the mean and standard deviation of the class so that grades are comparable across classes. We also include its square and cube. We further control for the percentile rank at the math and French DNB national exam, as well as for the percentile rank at the French oral and written *baccalauréat* anticipated examinations.
- *Previous year’s class characteristics:* It includes the average of all the students’ characteristics listed above computed at the Grade 11 level, the class average at the math continuous assessment, the lowest and the highest math grade of the class.

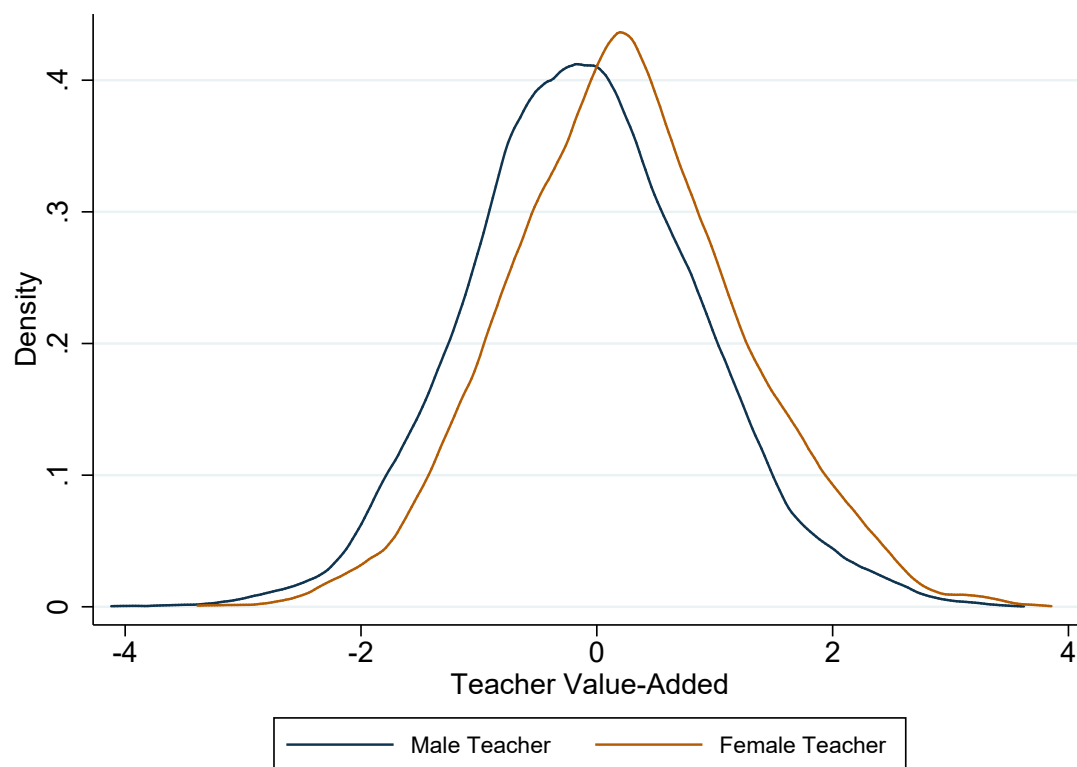
After the regression, we predict students' test scores residuals adjusted for observables.^{A.1} Finally, for each teacher's class in year t , we compute the average test score residual. This should be seen as a proxy for teachers quality in the class taught in year t .

Step 2: Regressing teachers' quality in year t on its lags and leads. We regress the average test score residuals of teachers in t on those average residuals in years $t - 1, t - 2, \dots$ and $t + 1, t + 2, \dots$. The OLS coefficients obtained from this regression tell us how strongly current teacher performance is related to its past and future performance, i.e. they are autocorrelation coefficients. These coefficients are also called *shrinkage* factors.

Step 3: Predicting teachers' quality. The final step consists in using the set of OLS coefficients from step 2 to *predict* teachers' quality. This predicted teacher quality is actually just a proxy for a teacher's true value-added and its reliability depends on the shrinkage factor, usually estimated to be around one-third (i.e. the true teacher value-added accounts for one-third of the residual variance).

The distribution of (standardized) teachers' predicted value-added is displayed in Figure 3.E1.

^{A.1}Teacher fixed effects are included in the regression so that coefficients on other covariates are estimated only using the within teacher variation. Those fixed effects are then added back to the residuals.

Figure 3.E1 – Distribution of Teachers' Predicted Value-Added

Notes: This graph plots the densities of math teachers' predicted value-added, separately for male and female teachers. The value-added estimates are obtained with the methodology described in Chetty et al. (2014) and implemented with the `vam` Stata package developed by Stepner (2013).

3.E.3 Complementary Results - Impact of Teacher GDV with Teacher Grading Bias or Value-Added Controls

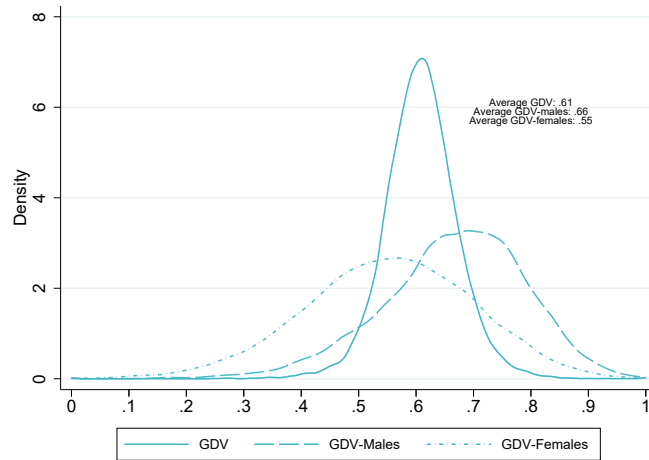
Table 3.E2 – Impact of Teacher GDV - Mechanisms - All outcomes

	Males			Females		
	Grading bias	Value added	Value added	Grading bias	Value added	Value added
	(1)	(2)	(3)	(4)	(5)	(6)
Academic performance						
Rank at <i>baccalauréat</i> : math	0.3714*** (0.0770)	0.3452*** (0.0763)	0.2486*** (0.0684)	0.3967*** (0.0851)	0.4084*** (0.0840)	0.2825*** (0.0727)
Rank at <i>baccalauréat</i> : philo	-0.0078 (0.0768)	-0.0039 (0.0763)	0.0033 (0.0766)	-0.0011 (0.0809)	0.0087 (0.0804)	0.0135 (0.0807)
Type of STEM programs ranked first in the ROL						
All STEM tracks	-0.0029*** (0.0011)	-0.0026** (0.0011)	-0.0027** (0.0011)	-0.0015 (0.0011)	-0.0015 (0.0010)	-0.0015 (0.0010)
Selective STEM	-0.0015 (0.0010)	-0.0013 (0.0010)	-0.0014 (0.0010)	0.0003 (0.0008)	0.0003 (0.0008)	0.0003 (0.0008)
<i>among which: biology</i>	-0.0003 (0.0003)	-0.0003 (0.0003)	-0.0003 (0.0003)	0.0000 (0.0004)	0.0001 (0.0004)	0.0001 (0.0004)
<i>among which: math, physics</i>	-0.0012 (0.0009)	-0.0010 (0.0009)	-0.0011 (0.0009)	0.0002 (0.0007)	0.0003 (0.0007)	0.0002 (0.0007)
University STEM	-0.0015*** (0.0006)	-0.0014** (0.0006)	-0.0014** (0.0006)	-0.0010 (0.0006)	-0.0010 (0.0006)	-0.0010 (0.0006)
Vocational STEM	0.0003 (0.0008)	0.0003 (0.0008)	0.0003 (0.0008)	-0.0010* (0.0005)	-0.0010* (0.0005)	-0.0010* (0.0005)
Matriculation in the following year						
All STEM	-0.0022** (0.0011)	-0.0021* (0.0011)	-0.0023** (0.0011)	-0.0014 (0.0010)	-0.0014 (0.0010)	-0.0014 (0.0010)
Selective STEM	-0.0025*** (0.0008)	-0.0024*** (0.0008)	-0.0025*** (0.0008)	-0.0004 (0.0007)	-0.0004 (0.0007)	-0.0005 (0.0007)
University STEM	0.0002 (0.0010)	0.0003 (0.0009)	0.0002 (0.0010)	-0.0010 (0.0008)	-0.0010 (0.0008)	-0.0010 (0.0008)
University Medicine	0.0007 (0.0007)	0.0007 (0.0006)	0.0007 (0.0006)	0.0002 (0.0009)	0.0002 (0.0009)	0.0002 (0.0009)
Vocational program	-0.0012*** (0.0004)	-0.0011*** (0.0004)	-0.0012*** (0.0004)	-0.0003 (0.0004)	-0.0002 (0.0004)	-0.0002 (0.0004)
Selective Other	0.0003 (0.0003)	0.0003 (0.0003)	0.0003 (0.0003)	0.0005 (0.0004)	0.0004 (0.0004)	0.0004 (0.0004)
University Other	0.0004 (0.0007)	0.0004 (0.0007)	0.0004 (0.0007)	-0.0001 (0.0008)	-0.0003 (0.0008)	-0.0004 (0.0008)
Not enrolled in Higher Ed	0.0020** (0.0008)	0.0018** (0.0008)	0.0020** (0.0008)	0.0008 (0.0009)	0.0010 (0.0009)	0.0012 (0.0009)
Nb. observations	340,729	342,380	342,380	300,020	301,409	301,409

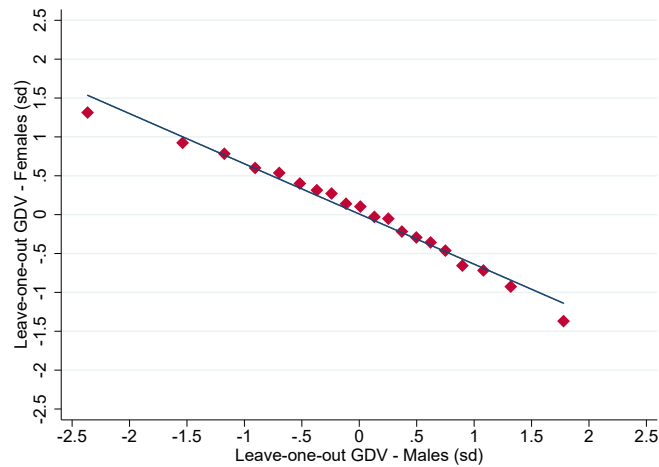
Notes:

3.E.4 GDV by Gender: Distribution and Correlation

Figure 3.E2 – Distribution and Correlation of Teachers Leave-one-out GDV by Gender



(a) Density of leave-one-out GDV



(b) Correlation

Panel (a) of this figure shows the distributions of math teachers' overall leave-one-out GDV, as well as the teacher accuracy computed for female students (*leave-one-out* GDV-females) and males students respectively (*leave-one-out* GDV-males). Panel (b) shows binned averages of GDV-males and GDV-females and plots the fitted regression line.

Bibliography

- Aakvik, Arild, Kjell G Salvanes, and Kjell Vaage**, “Measuring heterogeneity in the returns to education using an education reform,” *European Economic Review*, 2010, *54* (4), 483–500.
- Abdulkadiroğlu, Atila, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters**, “Do parents value school effectiveness?,” *American Economic Review*, 2020, *110* (5), 1502–39.
- Agarwal, Nikhil and Paulo Somaini**, “Demand analysis using strategic reports: An application to a school choice mechanism,” *Econometrica*, 2018, *86* (2), 391–444.
- Akerlof, George A and Rachel E Kranton**, “Economics and identity,” *The quarterly journal of economics*, 2000, *115* (3), 715–753.
- Akhtari, Mitra, Natalie Bau, and Jean-William P Laliberté**, “Affirmative Action and Pre-College Human Capital,” Technical Report, National Bureau of Economic Research 2020.
- Akyol, Pelin and Kala Krishna**, “Preferences, selection, and value added: A structural approach,” *European Economic Review*, 2017, *91*, 89–117.
- Alan, S., S. Ertac, and I. Mumcu**, “Gender Stereotypes in the Classroom and Effects on Achievement,” *Review of Economics and Statistics*, 2018, *100* (5), 876–890.
- Alfonsi, Livia, Oriana Bandiera, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali**, “Tackling youth unemployment: Evidence from a labor market experiment in Uganda,” *Econometrica*, 2020, *88* (6), 2369–2414.
- Altonji, Joseph G, Erica Blom, and Costas Meghir**, “Heterogeneity in human capital investments: High school curriculum, college major, and careers,” *Annu. Rev. Econ.*, 2012, *4* (1), 185–223.

- Andrews, Rodney J and Kevin M Stange**, “Price regulation, price discrimination, and equality of opportunity in higher education: evidence from Texas,” *American Economic Journal: Economic Policy*, 2019, 11 (4), 31–65.
- Angrist, Joshua D**, “The effect of veterans benefits on education and earnings,” *ILR Review*, 1993, 46 (4), 637–652.
- **and Alan B Krueger**, “Does compulsory school attendance affect schooling and earnings?,” *The Quarterly Journal of Economics*, 1991, 106 (4), 979–1014.
- Antonovics, Kate and Ben Backes**, “Were minority students discouraged from applying to University of California campuses after the affirmative action ban?,” *Education Finance and Policy*, 2013, 8 (2), 208–250.
- **and –**, “The effect of banning affirmative action on college admissions policies and student quality,” *Journal of Human Resources*, 2014, 49 (2), 295–322.
- Arcidiacono, Peter**, “Affirmative action in higher education: How do admission and financial aid rules affect future earnings?,” *Econometrica*, 2005, 73 (5), 1477–1524.
- **and Michael Lovenheim**, “Affirmative action and the quality-fit trade-off,” *Journal of Economic Literature*, 2016, 54 (1), 3–51.
- **, Josh Kinsler, and Tyler Ransom**, “Asian American Discrimination in Harvard Admissions,” Technical Report, National Bureau of Economic Research 2020.
- **, Michael Lovenheim, and Maria Zhu**, “Affirmative action in undergraduate education,” *Annu. Rev. Econ.*, 2015, 7 (1), 487–518.
- Ashworth, Jared, V Joseph Hotz, Arnaud Maurel, and Tyler Ransom**, “Changes across cohorts in wage returns to schooling and early work experiences,” 2017.
- Avery, Christopher N, Mark E Glickman, Caroline M Hoxby, and Andrew Metrick**, “A revealed preference ranking of us colleges and universities,” *The Quarterly Journal of Economics*, 2013, 128 (1), 425–467.

- Bagde, Surendrakumar, Dennis Epple, and Lowell Taylor**, “Does affirmative action work? Caste, gender, college quality, and academic success in India,” *American Economic Review*, 2016, *106* (6), 1495–1521.
- Bailey, Martha J and Susan M Dynarski**, “Gains and gaps: Changing inequality in US college entry and completion,” Technical Report, National Bureau of Economic Research 2011.
- Beilock, S., E. Gunderson, G. Ramirez, and S. Levine**, “Female Teachers’ Math Anxiety Affects Girls’ Math Achievement,” *Proceedings of the National Academy of Science*, 2010, *107* (5), 1860–1863.
- Bertrand, Marianne, Magne Mogstad, and Jack Mountjoy**, “Improving educational pathways to social mobility: Evidence from Norway’s “Reform 94”,” Technical Report, National Bureau of Economic Research 2019.
- Bettinger, Eric**, “How financial aid affects persistence,” in “College choices: The economics of where to go, when to go, and how to pay for it,” University of Chicago Press, 2004, pp. 207–238.
- Bettinger, Eric P, Bridget Terry Long, Philip Oreopoulos, and Lisa Sanbonmatsu**, “The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment,” *Quarterly Journal of Economics*, 2012, *127* (3), 1205–1242.
- Bhattacharya, Debopam, Shin Kanaya, and Margaret Stevens**, “Are university admissions academically fair?,” *Review of Economics and Statistics*, 2017, *99* (3), 449–464.
- Bian, L., S. Leslie, and A. Cimpian**, “Gender Stereotypes About Intellectual Ability Emerge Early and Influence Children’s Interests,” *Science*, 2017, *355*, 389–391.
- Bonneau, Cécile, Pauline Charousset, Julien Grenet, and Georgia Thebault**, *Quelle démocratisation des grandes écoles depuis le milieu des années 2000 ?*, Rapport IPP, 2021.
- Borhen, A., A. Imas, and M. Rosenberg**, “The Language of Discrimination: Using Experimental versus Observational Data,” *AEA Papers and Proceedings*, 2018, *108*, 169–174.
- Bound, John, Michael F Lovenheim, and Sarah Turner**, “Increasing time to baccalaureate degree in the United States,” *Education Finance and Policy*, 2012, *7* (4), 375–424.

- Bourdieu, Pierre and Monique de Saint Martin**, “Les catégories de l’entendement professoral,” *Actes de la recherche en sciences sociales*, 1975, 1 (3), 68–93.
- Breda, T., E. Jouini, and C. Napp**, “Societal Inequalities Amplify Gender Gaps in Math,” *Science*, 2018, 359 (6381).
- Breda, Thomas and Méлина Hillion**, “Teaching accreditation exams reveal grading biases favor women in male-dominated disciplines in France,” *Science*, 2016, 353 (6298), 474–478.
- **and Son Thierry Ly**, “Professors in core science fields are not always biased against women: Evidence from France,” *American Economic Journal: Applied Economics*, 2015, 7 (4), 53–75.
- **, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre**, “Do female role models reduce the gender gap in science? Evidence from classroom interventions in French high schools,” Technical Report, HAL 2020.
- Brodaty, Thomas, Robert J Gary-Bobo, and Ana Prieto**, “Does speed signal ability? The impact of grade repetitions on employment and wages,” *The Impact of Grade Repetitions on Employment and Wages (May 2008).*, Vol, 2008.
- Brown, Charles and Mary Corcoran**, “Sex-based differences in school content and the male-female wage gap,” *Journal of Labor Economics*, 1997, 15 (3), 431–465.
- Bucarey, Alonso**, “Who pays for free college? Crowding out on campus,” *Job market paper*, 2018.
- Cadinu, Mara, Anne Maass, Alessandra Rosabianca, and Jeff Kiesner**, “Why do women underperform under stereotype threat? Evidence for the role of negative thinking,” *Psychological science*, 2005, 16 (7), 572–578.
- Calsamiglia, Caterina, Guillaume Haeringer, and Flip Klijn**, “Constrained school choice: An experimental study,” *American Economic Review*, 2010, 100 (4), 1860–74.
- Card, David and Alan B Krueger**, “Would the elimination of affirmative action affect highly qualified minority applicants? Evidence from California and Texas,” *ILR Review*, 2005, 58 (3), 416–434.
- Carlana, M.**, “Implicit Stereotypes: Evidence from Teachers’ Gender Bias,” *Quarterly Journal of Economics*, 2019, 134 (3), 1163–1224.

- Carlana, Michela**, “Implicit stereotypes: Evidence from teachers’ gender bias,” *The Quarterly Journal of Economics*, 2019, *134* (3), 1163–1224.
- Carneiro, Pedro, James J Heckman, and Edward J Vytlacil**, “Estimating marginal returns to education,” *American Economic Review*, 2011, *101* (6), 2754–81.
- Castleman, Benjamin L and Lindsay C Page**, “Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates?,” *Journal of Economic Behavior & Organization*, 2015, *115*, 144–160.
- Chan, Jimmy and Erik Eyster**, “Does banning affirmative action lower college student quality?,” *American Economic Review*, 2003, *93* (3), 858–872.
- Cheng, A., , and Zamarro G. Kopotic K.**, “Can Parents’ Growth Mindset and Role Modelling Address STEM Gender Gaps?,” *Education Reform Faculty and Graduate Students Publications*, 2017.
- Cheryan, S. and V. Plaut**, “Explaining Underrepresentation: A Theory of Precluded Interest,” *Sex Roles*, 2010, *63*, 475–488.
- Chetty, R., J. Friedman, and J. Rockoff**, “Measuring the Impact of Teachers I: Evaluating Bias in Teacher Value-Added Estimates,” *American Economic Review*, 2014, *104* (9), 2593–2632.
- Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan**, “Mobility report cards: The role of colleges in intergenerational mobility,” Technical Report, national bureau of economic research 2017.
- Cnesco**, *La constitution des classes: pratiques et enjeux*, Paris: Cnesco, 2015.
- Copur-Gencturk, Y., I. Thacker, and D. Quinn**, “K-8 Teachers’ Overall and Gender-Specific Beliefs About Mathematical Aptitude,” *International Journal of Science and Mathematics Education*, 2020.
- Corpus, J. and M. Lepper**, “The Effects of Person Versus Performance Praise on Children’s Motivation: Gender and Age as Moderating Factors,” *Educational Psychology*, 2007, *27* (4), 487–508.

- Croizet, Jean-Claude, Gérard Després, Marie-Eve Gauzins, Pascal Huguet, Jacques-Philippe Leyens, and Alain Méot**, “Stereotype threat undermines intellectual performance by triggering a disruptive mental load,” *Personality and social psychology bulletin*, 2004, 30 (6), 721–731.
- Dalton, Patricio S, Sayantan Ghosal, and Anandi Mani**, “Poverty and aspirations failure,” *The Economic Journal*, 2016, 126 (590), 165–188.
- de Chaisemartin, Clément and Xavier d’Haultfoeuille**, “Two-way fixed effects estimators with heterogeneous treatment effects,” Technical Report, National Bureau of Economic Research 2019.
- **and Xavier d’Haultfoeuille**, “Fuzzy differences-in-differences,” *The Review of Economic Studies*, 2018, 85 (2), 999–1028.
- Demoulin, H. and C. Daniel**, “Bulletins scolaires et orientation au prisme du genre,” *L’orientation scolaire et professionnelle*, 2013, 42.
- DEPP**, “Le niveau d’études selon le niveau social,” 2020. Etat de l’Enseignement supérieur, de la recherche et de l’innovation, MENJS-MESRI-DEPP.
- Drewes, Torben and Christopher Michael**, “How do students choose a university?: an analysis of applications to universities in Ontario, Canada,” *Research in Higher Education*, 2006, 47 (7), 781–800.
- Duflo, Esther**, “Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment,” *American economic review*, 2001, 91 (4), 795–813.
- Dweck, C., W. Davidson, S. Nelson, and B. Enna**, “Sex Differences in Learned Helplessness: II. The Contingencies of Evaluative Feedback in the Classroom and III. An Experimental Analysis,” *Developmental Psychology*, 1978, 14 (3), 268–276.
- Dynarski, Susan, CJ Libassi, Katherine Micheltore, and Stephanie Owen**, “Closing the gap: The effect of a targeted, tuition-free promise on college choices of high-achieving, low-income students,” Technical Report, National Bureau of Economic Research 2018.

- Eble, Alex and Feng Hu**, “How Important are Beliefs about Gender Differences in Math Ability? Transmission across Generations and Impacts on Child Outcomes,” *EdWorkingPaper*, 2019, 19-67.
- and –, “Stereotypes, Role Models, and the Formation of Beliefs,” *CDEP-CGEG Working Paper No. 43*, 2019.
- Epple, Dennis, Richard Romano, and Holger Sieg**, “Diversity and affirmative action in higher education,” *Journal of Public Economic Theory*, 2008, 10 (4), 475–501.
- Fack, Gabrielle and Julien Grenet**, “Improving college access and success for low-income students: Evidence from a large need-based grant program,” *American Economic Journal: Applied Economics*, 2015, 7 (2), 1–34.
- , –, and **Yinghua He**, “Beyond truth-telling: Preference estimation with centralized school choice and college admissions,” *American Economic Review*, 2019, 109 (4), 1486–1529.
- Falch, Torberg, Astrid Marie Jorde Sandsør, and Bjarne Strøm**, “Do Smaller Classes Always Improve Students’ Long-run Outcomes?,” *Oxford Bulletin of Economics and Statistics*, 2017, 79 (5), 654–688.
- Field, Erica**, “Educational debt burden and career choice: Evidence from a financial aid experiment at NYU Law School,” *American Economic Journal: Applied Economics*, 2009, 1 (1), 1–21.
- Friedman, Jerome, Trevor Hastie, and Rob Tibshirani**, “Regularization Paths for Generalized Linear Models via Coordinate Descent,” *Journal of Statistical Software*, 2010, 33 (1), 1–22.
- Fu, Chao**, “Equilibrium tuition, applications, admissions, and enrollment in the college market,” *Journal of Political Economy*, 2014, 122 (2), 225–281.
- Gale, David E. and Lloyd S. Shapley**, “College Admissions and the Stability of Marriage,” *American Mathematical Monthly*, 1962, 69 (1), 9–15.
- Garibaldi, Pietro, Francesco Giavazzi, Andrea Ichino, and Enrico Rettore**, “College cost and time to complete a degree: Evidence from tuition discontinuities,” *Review of Economics and Statistics*, 2012, 94 (3), 699–711.

- Gentzkow, Matthew, Bryan Kelly, and Matt Taddy**, “Text as data,” *Journal of Economic Literature*, 2019, 57 (3), 535–74.
- Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift**, “Bartik instruments: What, when, why, and how,” *American Economic Review*, 2020, 110 (8), 2586–2624.
- Good, C., A. Rattan, and C. Dweck**, “Why Do Women Opt Out? Sense of Belonging and Women’s Representation in Mathematics,” *Journal of Personality and Social Psychology*, 2012, 102 (4), 700–717.
- Grenet, Julien**, “Is extending compulsory schooling alone enough to raise earnings? Evidence from French and British compulsory schooling laws,” *The Scandinavian Journal of Economics*, 2013, 115 (1), 176–210.
- Groen, Jeffrey A and Michelle J White**, “In-state versus out-of-state students: the divergence of interest between public universities and state governments,” *Journal of Public Economics*, 2004, 88 (9-10), 1793–1814.
- Guillois, P, C Lafargue, and O Monso**, “Les disparités territoriales dans la poursuite des études vers l’enseignement supérieur dans l’académie de Bordeaux,” *Géographie de l’école*, 2014, pp. 113–124.
- Guiso, L., F. Monte, P. Sapienza, and L. Zingales**, “Culture, Gender, and Math,” *Science*, 2008, 320.
- Gunderson, E., G. Ramirez, S. Levine, and S. Beilock**, “The Role of Parents and Teachers in the Development of Gender-Related Math Attitudes,” *Sex Roles*, 2012, 66, 153–166.
- Gurgand, Marc and Eric Maurin**, “Démocratisation de l’enseignement secondaire et inégalités salariales en France,” in “Annales. Histoire, sciences sociales,” Vol. 61 Éditions de l’EHESS 2006, pp. 845–859.
- Guyon, Nina and Elise Huillery**, “Biased aspirations and social inequality at school: Evidence from french teenagers,” *The Economic Journal*, 2016.
- and — , “Biased aspirations and social inequality at school: Evidence from french teenagers,” *The Economic Journal*, 2021, 131 (634), 745–796.

- , **Eric Maurin**, and **Sandra McNally**, “The effect of tracking students by ability into different schools a natural experiment,” *Journal of Human resources*, 2012, 47 (3), 684–721.
- Haeringer, Guillaume and Flip Klijn**, “Constrained school choice,” *Journal of Economic theory*, 2009, 144 (5), 1921–1947.
- and **Vincent Iehlé**, “Gradual college admission,” *Université Paris-Dauphine Research Paper*, 2019, (3488038).
- Hall, Caroline**, “Does more general education reduce the risk of future unemployment? Evidence from an expansion of vocational upper secondary education,” *Economics of Education Review*, 2016, 52, 251–271.
- Hampf, Franziska and Ludger Woessmann**, “Vocational vs. general education and employment over the life cycle: New evidence from PIAAC,” *CESifo Economic Studies*, 2017, 63 (3), 255–269.
- Hanushek, Eric A, Guido Schwerdt, Ludger Woessmann, and Lei Zhang**, “General education, vocational education, and labor-market outcomes over the lifecycle,” *Journal of human resources*, 2017, 52 (1), 48–87.
- Harmon, Colm and Ian Walker**, “Estimates of the economic return to schooling for the United Kingdom,” *The American Economic Review*, 1995, 85 (5), 1278–1286.
- Hastie, Trevor, Robert Tibshirani, and Jerome Friedman**, *The elements of statistical learning: data mining, inference, and prediction*, Springer Science & Business Media, 2009.
- Hastings, Justine, Christopher A Neilson, and Seth D Zimmerman**, “The effects of earnings disclosure on college enrollment decisions,” Technical Report, National Bureau of Economic Research 2015.
- Haveman, Robert and Barbara Wolfe**, “Succeeding generations: On the effects of investments in children,” 1994.
- Heckman, James J, John Eric Humphries, and Gregory Veramendi**, “Returns to education: The causal effects of education on earnings, health, and smoking,” *Journal of Political Economy*, 2018, 126 (S1), S197–S246.

- Hill, Francesca, Irene C Mammarella, Amy Devine, Sara Caviola, Maria Chiara Pas-solunghi, and Dénes Szűcs**, “Maths anxiety in primary and secondary school students: Gender differences, developmental changes and anxiety specificity,” *Learning and Individual Differences*, 2016, *48*, 45–53.
- Hoff, Karla and Priyanka Pandey**, “Discrimination, social identity, and durable inequalities,” *American economic review*, 2006, *96* (2), 206–211.
- Holzer, Harry and David Neumark**, “Assessing affirmative action,” *Journal of Economic literature*, 2000, *38* (3), 483–568.
- Howell, Jessica S**, “Assessing the impact of eliminating affirmative action in higher education,” *Journal of Labor Economics*, 2010, *28* (1), 113–166.
- Hoxby, Caroline M and Sarah Turner**, “What high-achieving low-income students know about college,” *American Economic Review*, 2015, *105* (5), 514–17.
- Jacob, Brian, Brian McCall, and Kevin Stange**, “College as country club: Do colleges cater to students’ preferences for consumption?,” *Journal of Labor Economics*, 2018, *36* (2), 309–348.
- Jaquette, Ozan and Bradley R Curs**, “Creating the out-of-state university: Do public univer-sities increase nonresident freshman enrollment in response to declining state appropriations?,” *Research in Higher Education*, 2015, *56* (6), 535–565.
- Jensen, Robert**, “The (perceived) returns to education and the demand for schooling,” *Quarterly Journal of Economics*, 2010, *125* (2), 515–548.
- Kane, Thomas J**, “Racial and ethnic preferences in college admissions,” *Ohio St. LJ*, 1998, *59*, 971.
- Kapor, Adam**, “Distributional effects of race-blind affirmative action,” *Job Market Paper, Yale University*, 2015.
- Keller, C.**, “Effect of Teachers’ Stereotyping on Students’ Stereotyping of Mathematics as a Male Domain,” *The Journal of Social Psychology*, 2001, *141* (2), 165–173.
- Kirkebøen, Lars Johannessen**, “Preferences for lifetime earnings, earnings risk and nonpecuniary attributes in choice of higher education,” Technical Report, Discussion Papers 2012.

- Koffi, M.**, “Innovative Ideas and Gender Inequality,” *Job Market Paper*, 2020.
- Kojima, Fuhito**, “School choice: Impossibilities for affirmative action,” *Games and Economic Behavior*, 2012, *75* (2), 685–693.
- Lavy, Victor and Edith Sand**, “On the origins of gender gaps in human capital: Short- and long-term consequences of teachers’ biases,” *Journal of Public Economics*, 2018, *167* (C), 263–279.
- **and —**, “On the origins of gender gaps in human capital: Short-and long-term consequences of teachers’ biases,” *Journal of Public Economics*, 2018, *167*, 263–279.
- Leslie, S., A. Cimpian, M. Meyer, and E. Freeland**, “Expectations of Brilliance Underlie Gender Distribution Across Academic Disciplines,” *Science*, 2015, *347* (6219).
- Leuven, Edwin and Sturla A Løkken**, “Long-term impacts of class size in compulsory school,” *Journal of Human Resources*, 2020, *55* (1), 309–348.
- , **Hessel Oosterbeek, and Marte Rønning**, “Quasi-experimental estimates of the effect of class size on achievement in Norway,” *Scandinavian Journal of Economics*, 2008, *110* (4), 663–693.
- Lim, Jaegeum and Jonathan Meer**, “The Impact of Teacher-Student Gender Matches: Random Assignment Evidence from South Korea,” *Journal of Human Resources*, 2017, pp. 1215–7585.
- Lindberg, S., J. Hyde, J. Peterson, and M. Linn**, “New Trends in Gender and Mathematics Performance: A Meta-Analysis,” *Psychological Bulletin*, 2010, *136* (6), 1123–1136.
- Lochner, Lance and Enrico Moretti**, “The effect of education on crime: Evidence from prison inmates, arrests, and self-reports,” *American economic review*, 2004, *94* (1), 155–189.
- Long, Mark C**, “Race and college admissions: An alternative to affirmative action?,” *review of Economics and Statistics*, 2004, *86* (4), 1020–1033.
- **and Marta Tienda**, “Winners and losers: Changes in Texas university admissions post-Hopwood,” *Educational Evaluation and Policy Analysis*, 2008, *30* (3), 255–280.
- Luflade, Margaux**, “The value of information in centralized school choice systems (job market paper),” 2017.

- Ly, Son Thierry and Arnaud Riegert**, “Mixité sociale et scolaire et ségrégation inter—et intra-établissement dans les collèges et lycées français,” *Rapport du Conseil national d’évaluation du système scolaire (CNESEO)*. Téléaccessible à: <http://www.cnesco.fr/wp-content/uploads/2015/05/Etat-des-lieux-Mixité-à-lécoleFrance1.pdf>, 2015.
- Machin, Stephen J, Sandra McNally, Camille Terrier, and Guglielmo Ventura**, “Closing the Gap between Vocational and General Education? Evidence from University Technical Colleges in England,” 2020.
- Malamud, Ofer and Cristian Pop-Eleches**, “General education versus vocational training: Evidence from an economy in transition,” *The review of economics and statistics*, 2010, *92* (1), 43–60.
- Mane, Ferran**, “Trends in the payoff to academic and occupation-specific skills: the short and medium run returns to academic and vocational high school courses for non-college-bound students,” *Economics of education review*, 1999, *18* (4), 417–437.
- Maurin, Eric and Sandra McNally**, “Vive la revolution! Long-term educational returns of 1968 to the angry students,” *Journal of Labor Economics*, 2008, *26* (1), 1–33.
- McFadden, Daniel et al.**, “Conditional logit analysis of qualitative choice behavior,” 1973.
- Meer, Jonathan**, “Evidence on the returns to secondary vocational education,” *Economics of education review*, 2007, *26* (5), 559–573.
- Meghir, Costas and Mårten Palme**, “Educational reform, ability, and family background,” *American Economic Review*, 2005, *95* (1), 414–424.
- MENESR**, *L’état de l’Enseignement supérieur et de la Recherche en France - Édition 2017*, Ministère de l’Éducation nationale, de l’Enseignement supérieur et de la Recherche, 2017.
- Miller, D., A. Eagly, and M. Linn**, “Women’s Representation in Science Predicts National Gender-Science Stereotypes: Evidence From 66 Nations,” *Journal of Educational Psychology*, 2014, *107*, 631–644.
- Morgan, C.**, “The Effect of Negative Managerial Feedback on Student Motivation: Implications for Gender Differences in Teacher-Student Relations,” *Sex Roles*, 2001, *44*.

- Ningrum, P., T. Pansombut, and A. Ueranantasun**, “Text Mining of Online Job Advertisements to Identify Direct Discrimination During Job Hunting Process: A Case Study in Indonesia,” *PLoS ONE*, 2020, 15 (6).
- Nosek, Brian A., Frederick L. Smyth, N. Sriram, Nicole M. Lindner, Thierry Devos, Alfonso Ayala, Yoav Bar-Anan, Robin Bergh, Huajian Cai, Karen Gonsalkorale, Selin Kesebir, Norbert Maliszewski, Félix Neto, Eero Olli, Jaihyun Park, Konrad Schnabel, Kimihiro Shiomura, Bogdan Tudor Tulbure, Reinout W. Wiers, Mónica Somogyi, Nazar Akrami, Bo Ekehammar, Michelangelo Vianello, Mahzarin R. Banaji, and Anthony G. Greenwald**, “National Differences in Gender–Science Stereotypes Predict National Sex Differences in Science and Math Achievement,” 2009, 106 (26), 10593–10597.
- Nosek, Brian A, Frederick L Smyth, Natarajan Sriram, Nicole M Lindner, Thierry Devos, Alfonso Ayala, Yoav Bar-Anan, Robin Bergh, Huajian Cai, Karen Gonsalkorale et al.**, “National differences in gender–science stereotypes predict national sex differences in science and math achievement,” *Proceedings of the National Academy of Sciences*, 2009, 106 (26), 10593–10597.
- Oguzoglu, U. and S. Ozbeklik**, “Like Father, Like Daughter (Unless There Is a Son): Sibling Sex Composition and Women’s STEM Major Choice in College,” *IZA Discussion Papers*, 2016, 10052.
- Oosterbeek, Hessel and Dinand Webbink**, “Wage effects of an extra year of basic vocational education,” *Economics of Education Review*, 2007, 26 (4), 408–419.
- Oreopoulos, Philip**, “Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling,” *Journal of public Economics*, 2007, 91 (11-12), 2213–2229.
- **and Ryan Dunn**, “Information and college access: Evidence from a randomized field experiment,” *The Scandinavian Journal of Economics*, 2013, 115 (1), 3–26.
- Oreopoulos, Philipp and Uros Petronijevic**, “Making College Worth it: A Review of the Returns to Higher Education,” *The Future of Children*, 2013, 23 (1).

- Ost, Ben, Weixiang Pan, and Douglas Webber**, “The returns to college persistence for marginal students: Regression discontinuity evidence from university dismissal policies,” *Journal of Labor Economics*, 2018, *36* (3), 779–805.
- Oyserman, Daphna, Deborah Bybee, and Kathy Terry**, “Possible selves and academic outcomes: How and when possible selves impel action.,” *Journal of personality and social psychology*, 2006, *91* (1), 188.
- Piketty, Thomas, Mathieu Valdenaire et al.**, *L’impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français: estimations à partir du panel primaire 1997 et du panel secondaire 1995*, Direction de l’évaluation et de la prospective, 2006.
- Porter, Catherine and Danila Serra**, “Gender differences in the choice of major: The importance of female role models,” *American Economic Journal: Applied Economics*, 2020, *12* (3), 226–54.
- Reid, Stephen and Rob Tibshirani**, “Regularization Paths for Conditional Logistic Regression: The clogitL1 Package,” *Journal of Statistical Software*, 2014, *58* (12), 1–23.
- Roth, Alvin E**, “The economics of matching: Stability and incentives,” *Mathematics of operations research*, 1982, *7* (4), 617–628.
- Saltiel, Fernando**, “Fast-Tracked to Success: Evidence on the Returns to Vocational Education in Switzerland,” *Annals of Economics and Statistics (forthcoming)*, 2020.
- Sander, Richard H**, “A systemic analysis of affirmative action in American law schools,” *Stan. L. Rev.*, 2004, *57*, 367.
- Skaalvik, Sidsel and Einar M Skaalvik**, “Gender differences in math and verbal self-concept, performance expectations, and motivation,” *Sex roles*, 2004, *50* (3), 241–252.
- Skočajić, M., Radosavljević J., M. Okičić, I. Janković, and I. Žeželj**, “Boys Just Don’t! Gender-Stereotyping and Sacntionning of Counterstereotypical Behavior in Preschoolers,” *Sex Roles*, 2020, *82*, 163–172.
- Smith, Jonathan, Matea Pender, and Jessica Howell**, “The full extent of student-college academic undermatch,” *Economics of Education Review*, 2013, *32*, 247–261.

-
- Stepner, M.**, “VAM: Stata Module to Compute Teacher Value-Added Measures,” *Statistical Software Components*, 2013, S457711.
- Stone, Jeff, Christian I Lynch, Mike Sjomeling, and John M Darley**, “Stereotype threat effects on black and white athletic performance.,” *Journal of personality and social psychology*, 1999, 77 (6), 1213.
- Terrier, C.**, “Boys Lag Behind: How Teachers’ Gender Biases Affect Student Achievement,” *Economics of Education Review*, 2020, 77.
- Thomas, James**, “What do course offerings imply about university preferences?,” *FTC Bureau of Economics Working Paper*, 2019, (340).
- Wu, A.**, “Gendered Language on the Economics Job Market Rumors Forum,” *AEA Papers and Proceedings*, 2018, 108, 175–179.
- Wu, X., C. Li, Y. Zhu, and Y. Miao**, “Short Text Topic Modeling with Topic Distribution Quantization and Negative Sampling Decoder,” *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*, 2020.
- Zanten, Agnès Van**, “5. Les inégalités d’accès à l’enseignement supérieur,” *Regards croisés sur l’économie*, 2015, (1), 80–92.
- Zilic, Ivan**, “General versus vocational education: Lessons from a quasi-experiment in Croatia,” *Economics of Education Review*, 2018, 62, 1–11.
- Zimmerman, Seth D**, “The returns to college admission for academically marginal students,” *Journal of Labor Economics*, 2014, 32 (4), 711–754.

List of Tables

1.1	Odds Ratios of the Probability of BEP Program Conversion in Year 2009, 2010, or 2011 (base 2008)	68
1.2	Pre-Reform Shares Assigned by Middle Schools to High Schools	69
1.3	Summary Statistics: Vocational vs. Non-Vocational Students	69
1.4	Descriptive Evidence on Reform’s Impact: Vocational vs. non-Vocational Students	70
1.5	Reform’s Reduced-Form Impact: Educational Outcomes	71
1.6	Reform’s Reduced-Form Impact: Educational and Labor Market Outcomes - Wage Earners Only	72
1.A1	Statistical Match Between Educational and Labor Market Data	77
1.A2	Reform’s Reduced-Form Impact on Vocational Students’ Characteristics	79
1.C1	Reform’s Reduced-Form Impact on Educational and Labor Market Outcomes, by Gender - Wage Earners Only	84
1.C2	Reform’s Reduced-Form Impact on Educational and Labor Market Outcomes, by SES - Wage Earners Only	85
1.C3	Reform’s Reduced-Form Impact on Educational and Labor Market Outcomes, by Prior Achievement Level - Wage Earners Only	86
1.D1	Reform’s Reduced-Form Impact - Wald-TC Estimator	89
1.D2	Reform’s Reduced-Form Impact - Vocational and Non-Vocational Students, Wage Earners Only	91
1.D3	Reform’s Reduced-Form Impact Using Variation at the <i>Département</i> Level - Wage Earners Only	93
1.D4	Reform’s Reduced-Form Impact Using Variation at the <i>Académie</i> Level- Wage Earners Only	94
1.E1	Reform’s Potential Mechanisms - Wage Earners Only	96

1.E2 Reform’s Reduced-Form Impact on Vocational Students’ Outcomes - Wage Earners Only, <i>Départements</i> without CAP effect	98
1.E3 Reform’s Reduced-Form Impact on Vocational Students’ Outcomes - Wage Earners Only, <i>Départements</i> without Class Size Adjustment	99
1.E4 Reform’s Reduced-Form Impact on Vocational Students’ Outcomes, by Predicted Pre-Reform Agent Type	102
1.E5 Male-Dominated vs. Female-Dominated Vocational Specialisations	103
1.E6 Wages and Employment Rates Nine Years After Grade 9: Male-Dominated vs. Female-Dominated Vocational Specializations	104
2.1 Summary Statistics: Applicants to CPGE Programs	143
2.2 Summary Statistics: CPGE vs. Other Post-Secondary Programs	144
2.3 Summary Statistics: Ranked vs. Unranked Applications in CPGE	145
2.4 Summary Statistics: Applications by Quartile of Program’s Ranking	146
2.5 Reverse Engineered Weights on Applicants’ Characteristics	147
2.6 Characteristics of CPGE Programs Assigning Bottom Quartile Weight to Applicants’ Characteristics	148
2.7 Characteristics of CPGE Programs Assigning Top Quartile Weight to Applicants’ Characteristics	149
2.8 Policy Ban on Applicants’ Characteristics: Impact on CPGE Composition	150
2.9 Policy Ban on Applicants’ Characteristics: Impact on Matching Outcomes	151
2.A1 Replication of the Gale Shapley Deferred Acceptance Algorithm	156
2.B1 Summary Statistics: Applicants to Science CPGE	157
2.B2 Summary Statistics: Applicants to Business CPGE	158
2.B3 Summary Statistics: Applicants to Humanities CPGE	159
2.D1 Reverse Engineered Weights on Grades	166
2.D2 Counterfactual Assignment Simulations: CPGE Composition by Quartile of Selec- tivity	167
2.D3 Counterfactual Assignment Simulations: CPGE Composition by Geographical Loca- tion	168
2.E1 Correlation Between Observed and Simulated Baseline Rankings of Applicants	169
2.F1 Share Students with Non-Missing Values on Grades	171

2.G1 Predicted Probability of Access to a Top-Tier Grande École	172
3.1 Number of Grade 12 Students and Sample Restrictions	217
3.2 Students' Summary Statistics	218
3.3 Teachers' Summary Statistics	219
3.4 Balancing Test: Teachers' Leave-One-Out GDV with Students' Baseline Characteristics	220
3.5 Pearson's Chi Square Tests of Class Random Assignment	221
3.6 Impact of Teacher GDV on Academic Performance, Higher Education Choices and Enrollment	222
3.7 Impact of Teacher GDV - Mechanisms	223
3.8 Impact of Teacher GDV-Females and GDV-Males - Mechanisms	224
3.A1 From Text to Data: An Illustration	228
3.A2 Top 100 Predictors' Classification - Female	231
3.A3 Top 100 Predictors' Classification - Male	232
3.B1 Share of Male Teachers by Core Subjects	233
3.C1 Balancing Test: High Schools With All Missing Grade Transcripts	234
3.D1 Impact of Teacher GDV - Robustness Checks	236
3.E1 Teachers' Average Math Grading Bias for Grade 12 Students	242
3.E2 Impact of Teacher GDV - Mechanisms - All outcomes	246

List of Figures

1.1	Structure of the French Secondary Education System	55
1.2	Vocational Students' Educational Pathways Before the Reform	56
1.3	BEP Conversions and Vocational Students' Enrollment Status After Grade 9	56
1.4	Graphical Evidence on the Impact of the Reform on Student Outcomes in Secondary Education: Vocational vs. Non-Vocational Students	57
1.5	Graphical Evidence on the Impact of the Reform on Student Outcomes in Higher Education: Vocational vs. Non-Vocational Students	58
1.6	Graphical Evidence on the Impact of the Reform on Labor Market Outcomes Nine Years after Grade 9: Vocational vs. Non-Vocational Students	59
1.7	Illustration of Empirical Strategy: Secondary Education Outcomes by Predicted Exposure to Early Vocational Program Conversions	60
1.8	Illustration of Empirical Strategy: Higher Education Outcomes by Predicted Exposure to Early Vocational Program Conversions	61
1.9	Illustration of Empirical Strategy: Labor Market Outcomes Nine Years after Grade 9 by Predicted Exposure to Early Vocational Program Conversions	62
1.10	Reduced-Form Coefficients on Exposure to Vocational Program Conversions	63
1.11	Reduced-Form Coefficients on Exposure to Vocational Program Conversions: Female vs. Male Students	64
1.12	Reduced-Form Coefficients on Exposure to Vocational Program Conversions: Low vs. Medium-Low to High SES	65
1.13	Reduced-Form Coefficients on Exposure to Vocational Program Conversions: Low vs. High Prior Achievement Level	66
1.14	Coefficients from the Quantile Regression of Log Annual Wages on Exposure to the Converted Programs	67

1.A1 Student Tracking after Grade 9	78
1.B1 Reform’s Reduced-Form Impact on Grade Retention and High School Dropout	83
2.1 Estimated Weights on Applicants’ Grades: Science CPGE	134
2.2 Estimated Weights on Applicants’ Grades: Business CPGE	135
2.3 Estimated Weights on Applicants’ Grades: Humanities CPGE	136
2.4 Estimated Weights on Applicants’ Demographic Characteristics: Science CPGE	137
2.5 Estimated Weights on Applicants’ Demographic Characteristics: Business CPGE	138
2.6 Estimated Weights on Applicants’ Demographic Characteristics: Humanities CPGE	139
2.7 Applicants’ Characteristics by Decile of Programs’ Rankings: Science CPGE	140
2.8 Applicants’ Characteristics by Decile of Programs’ Rankings: Business CPGE	141
2.9 Applicants’ Characteristics by Decile of Programs’ Rankings: Humanities CPGE	142
3.1 Share of Male and Female Students by Quartile of Percentile Rank at the DNB Math Exam	205
3.2 Enrollment in Higher Education by Gender after Grade 12 Science Track	206
3.3 Math Feedback - Distribution of Word Counts	206
3.4 Distribution of Teachers’ Gender Gaps in Feedback Type and Positiveness - Females minus Males	207
3.5 Distribution of Math Teachers GDV and Leave-one-out GDV	208
3.6 Distribution of Math Teachers’ Leave-one-out GDV – By Teacher Gender	209
3.7 Distribution of Teachers’ Leave-one-out GDV – By Core Subjects	210
3.8 Odds Ratios of the Top 10 Gender Predictors	211
3.9 Classification of the Top 30 Gender Predictors	212
3.10 Gender Predictors’ Type and Positiveness	213
3.11 Teachers’ Gender Gap in the Share of Positive Words in Favor of Females by Deciles of GDV - In Absolute Value and Percentage	213
3.12 Effect of Math Teacher GDV on Math Performance at <i>Baccalauréat</i> - By GDV Deciles	214
3.13 Effect of Math Teacher GDV on STEM Programs First Choice	215
3.14 Effect of Math Teacher GDV on Matriculation in the Following Year	216
3.15 Correlation Between Teacher GDV, Grading Bias and Teacher Quality	216
3.D1 Impact of Teacher GDV on Math Performance at <i>Baccalauréat</i> - By Deciles of Initial Math Performance	238

3.D2 Impact of Teacher GDV on Math Performance at <i>Baccalauréat</i> - By Social Background	240
3.E1 Distribution of Teachers' Predicted Value-Added	244
3.E2 Distribution and Correlation of Teachers Leave-one-out GDV by Gender	249