

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

L'EFFET DE TRAITEMENT DES ÉCOLES PRIVÉES AU QUÉBEC SUR  
LES SCORES AUX TESTS PISA DE 2000 À 2012 : UNE ANALYSE PAR  
MÉTHODE D'APPARIEMENT

MÉMOIRE  
PRÉSENTÉ  
COMME EXIGENCE PARTIELLE  
DE LA MAÎTRISE EN ÉCONOMIQUE

PAR  
DAVID LAPIERRE

MAI 2016

UNIVERSITÉ DU QUÉBEC À MONTRÉAL  
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.07-2011). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

## REMERCIEMENTS

« La présente recherche a été menée grâce à un soutien financier accordé au Réseau canadien des centres de données de recherche (RCCDR), par le Conseil de recherches en sciences humaines (CRSH), les Instituts de recherche en santé du Canada (IRSC), la Fondation canadienne pour l'innovation (FCI) et Statistique Canada. Bien que les recherches et les analyses aient été faites à partir des données de Statistique Canada, les opinions exprimées ne représentent pas celles de Statistique Canada. »

Ce mémoire de maîtrise a pu être réalisé grâce aux bourses de recherche accordées par le Groupe de recherche sur le capital humain et des professeurs Pierre Lefebvre et Philip Merrigan, ainsi que celles offertes par le Centre interuniversitaire québécois de statistique sociale. Je remercie les professeurs, chercheurs et membres du personnel de ces institutions pour leur travail et leur aide.

## TABLE DES MATIÈRES

REMERCIEMENTS . . . . .	ii
LISTE DES TABLEAUX . . . . .	v
RÉSUMÉ . . . . .	vii
INTRODUCTION . . . . .	1
CHAPITRE I	
REVUE DE LITTÉRATURE . . . . .	5
1.1 Revue des études sur l'école privée . . . . .	5
1.1.1 Impact de l'école privée sur ses élèves . . . . .	5
1.1.2 Effet d'« écrémage » et effets de pairs . . . . .	9
1.2 Revue de méthodologie économétrique . . . . .	11
1.2.1 Balancement entropique . . . . .	11
1.2.2 Choix du modèle . . . . .	12
1.2.3 Sensibilité aux variables confondantes . . . . .	13
CHAPITRE II	
DONNÉES ET VARIABLES . . . . .	15
2.1 Présentation des données . . . . .	15
2.2 Valeurs plausibles . . . . .	17
2.3 Cohérence des données . . . . .	17
2.4 Variables et statistiques descriptives . . . . .	19
CHAPITRE III	
MÉTHODOLOGIE . . . . .	27
3.1 Méthode échantillonnale . . . . .	27
3.2 Poids d'échantillonnage et poids répliqués . . . . .	28
3.2.1 Poids d'échantillonnage . . . . .	28
3.2.2 Poids répliqués . . . . .	29

3.3	Régression logistique . . . . .	30	
3.4	Balancement entropique . . . . .	31	
3.5	Régressions par moindres carrés . . . . .	32	
3.6	Effets de pairs . . . . .	34	
3.7	Simulation de variable confondante . . . . .	34	
CHAPITRE IV			
PRÉSENTATION DES RÉSULTATS . . . . .			37
4.1	Estimations . . . . .	37	
4.1.1	Régression logistique . . . . .	37	
4.1.2	Estimations générales . . . . .	39	
4.1.3	Estimations par sexe . . . . .	39	
4.2	Simulation de variable confondante . . . . .	40	
CHAPITRE V			
DISCUSSION DES RÉSULTATS . . . . .			43
CONCLUSION . . . . .			45
APPENDICE A			
TABLEAUX SUPPLÉMENTAIRES . . . . .			47
BIBLIOGRAPHIE . . . . .			62

## LISTE DES TABLEAUX

Tableau	Page
2.1 Type d'établissement par année . . . . .	19
2.2 Scores PISA : public et privé . . . . .	20
2.3 Scores PISA : filles et garçons . . . . .	21
2.4 Quintile de statut socio-économique selon le type d'établissement	23
2.5 Éducation de la mère en 2000 selon le type d'établissement . . . .	24
2.6 Éducation de la mère post-2000 selon le type d'établissement . . .	25
3.1 Probabilités de sélection d'école, intra-école et finale et poids asso- ciés pour un échantillon PPT avec écoles de tailles inégales . . . . .	28
4.1 Régression logistique - probabilité de fréquenter une école privée .	38
4.2 Effet de traitement par matière . . . . .	39
4.3 Effet de traitement par matière : filles . . . . .	40
4.4 Effet de traitement par matière : garçons . . . . .	40
4.5 Simulation de variable confondante : mathématiques . . . . .	41
4.6 Simulation de variable confondante : lecture . . . . .	41
4.7 Simulation de variable confondante : science . . . . .	42
5.1 Scores PISA : écarts entre public et privé (pourcentages d'écart-type)	43
A.1 Statistiques générales, 1/3 . . . . .	48
A.2 Statistiques générales, 2/3 . . . . .	49
A.3 Statistiques générales, 3/3 . . . . .	50
A.4 Répartition des sexes selon le type d'établissement . . . . .	51

A.5	Types d'établissement selon la taille des villes . . . . .	51
A.6	Nombre de parents immigrants selon le type d'établissement . . .	52
A.7	Langue parlée à la maison selon le type d'établissement . . . . .	52
A.8	Score en mathématiques . . . . .	53
A.9	Score en lecture . . . . .	54
A.10	Score en science . . . . .	55
A.11	Score des filles en mathématiques . . . . .	56
A.12	Score des filles en lecture . . . . .	57
A.13	Score des filles en science . . . . .	58
A.14	Score des garçons en mathématiques . . . . .	59
A.15	Score des garçons en lecture . . . . .	60
A.16	Score des garçons en science . . . . .	61

## RÉSUMÉ

On observe au Québec un écart de performance considérable aux tests nationaux et internationaux entre les étudiants des écoles privées subventionnées et ceux fréquentant le système public. Celui-ci s'élève à plus d'une année d'enseignement selon les résultats de l'enquête du *Programme for International Student Assessment* (PISA) réalisée en 2012 par l'OCDE.

Cet écart de performance ne peut toutefois pas être exclusivement attribué au type d'établissement fréquenté par l'élève. Les biais de sélection, de recrutement ou de causalité proscrivent les comparaisons simplistes. Par exemple, le statut socio-économique de l'individu aura un impact sur le choix de l'établissement d'enseignement, mais également sur sa performance académique.

Ce mémoire de maîtrise porte sur l'effet de traitement de l'école privée, en utilisant la méthode du balancement entropique comme solution à cette problématique. Cette méthode consiste à modifier les poids d'échantillonnage afin d'équilibrer les moments des variables choisies par le chercheur afin de créer un contre-factuel statistique au groupe traité, permettant au chercheur d'effectuer des estimations non-biaisées qui permettront d'isoler l'effet du type d'établissement. La simulation de variable confondante est également employée afin de vérifier si les résultats sont robustes à l'omission potentielle d'une variable corrélée à la fois avec le choix du traitement et un résultat positif, nonobstant l'assignation au traitement.

Les résultats obtenus sont importants et significatifs à 99 %. On observe en effet un effet de traitement allant de 29 % à 34 % d'un écart-type, soit plus de 53 % de l'écart initialement observé dans les données. Ces résultats sont également robustes à la simulation de variable confondante.

**MOTS-CLÉS** : éducation, école privée, capital humain, Québec, effet de traitement, méthodes d'appariement, balancement entropique, simulation de variable confondante.

## INTRODUCTION

En 2000, l'Organisation de coopération et de développement économique (OCDE) lançait l'étude PISA (*Programme for International Student Assessment*), cherchant à évaluer et comparer les niveaux d'habiletés cognitives des étudiants de 15 ans des pays participants dans une optique d'amélioration des politiques et des résultats. Ce faisant, on reconnaissait l'impact de la qualité du système d'éducation sur le niveau de développement des habiletés cognitives des jeunes et ses effets potentiels sur la réussite académique, le développement du capital humain de la population et ses conséquences sur l'innovation, la productivité et la croissance économique. La recherche à ce sujet est exhaustive et démontre le lien entre le développement des habiletés cognitives ainsi qu'académiques et le statut socioéconomique de l'individu adulte (Hanushek et Woessmann, 2008; Murnane *et al.*, 1995; Murnane *et al.*, 2000).

Toutefois, la relation entre les dépenses gouvernementales en éducation et les résultats obtenus par les élèves est difficile à établir et n'est certainement pas linéaire (Vandenberghe et Robin, 2004). Ceci soulève la question de l'organisation du système d'éducation, c'est-à-dire qu'il faut déterminer si différentes méthodes d'organisation et d'administration des écoles peuvent fournir des résultats supérieurs à ceux obtenus par un système public traditionnel.

La situation du Québec nous apparaît comme un cas intéressant à étudier en raison du taux de fréquentation très élevé des écoles privées par les élèves du secondaire. En effet, jusqu'à 12 % des élèves québécois fréquentent un établissement privé (ce chiffre s'élevant à 20 % pour les élèves du secondaire), alors que la moyenne

canadienne ne s'établit qu'à 6 %, ce qui s'explique par les subventions relativement généreuses offertes par le gouvernement provincial, couvrant au moins la moitié des frais scolaires encourus par les parents (la subvention s'élevant à 60 % du montant dépensé par le gouvernement pour un élève dans le système public, et l'établissement ne peut exiger davantage que ce même montant aux parents et facture généralement moins).

Lorsque sont observés les scores aux tests internationaux (PISA, TIMMS, PIRLS), il appert que les élèves québécois performant relativement bien dans les domaines étudiés (lecture, mathématiques, sciences). Toutefois, en comparant les résultats des tests PISA 2012 des élèves québécois en mathématiques selon le type d'établissement fréquenté (public ou privé), on constate une différence de soixante points entre les moyennes des scores des deux groupes. Cela équivaldrait à un écart d'un niveau de compétence entier entre les deux catégories d'élèves selon l'évaluation de l'OCDE, qui comporte six niveaux de compétence. Les élèves des établissements publics seraient d'un niveau 3 en moyenne (65,2 % des élèves canadiens performeraient à ce niveau ou à un niveau supérieur), alors que les élèves fréquentant les établissements privés atteindraient en moyenne le niveau 4 (atteint ou surpassé par 38,8 % des élèves canadiens). Un écart de quarante points correspondrait à une année d'enseignement entière. Selon l'OCDE, le niveau 2 serait le niveau minimal de compétence en mathématiques requis afin de pouvoir participer à la société moderne. L'effet de l'éducation au privé sur les résultats en sciences et en lecture est moindre mais demeure significatif.

Il est toutefois inapproprié d'attribuer exclusivement ces différences de résultats au type d'école fréquentée par l'élève. Le problème réside dans la possibilité pour les écoles privées de sélectionner les élèves sur la base de leurs tests d'aptitude, ainsi que dans l'auto-sélection des élèves selon la capacité de payer de leurs parents, alors que la plupart des écoles publiques n'ont pas la liberté d'effectuer cette

sélection. Cette différence cause des problèmes méthodologiques, tels que le biais de sélection, de causalité ou de recrutement. Par exemple, le niveau d'éducation des parents aura un impact sur la valorisation de l'éducation par ceux-ci ; de même qu'un statut socio-économique supérieur leur permettra d'accorder des sommes plus importantes à l'achat d'intrants éducatifs pouvant augmenter le niveau de leurs habiletés cognitives. Ces facteurs influencent ainsi à la fois le choix du type d'école, de même que les résultats scolaires de l'enfant en raison des ressources plus importantes mises à sa disposition qui favoriseront sa réussite scolaire.

Nous proposons l'utilisation de la méthode d'appariement dite de balancement entropique afin d'éliminer ou de réduire les effets de milieu liés à l'environnement socioéconomique et d'isoler l'impact de l'école privée sur les résultats aux tests PISA en mathématiques, en sciences et en lecture des étudiants québécois pour les années de 2000 allant jusqu'en 2012. Cette méthode, développée par Hainmueller (Hainmueller, 2012), offre une solution intéressante aux problèmes de biais et d'appariement tout en conservant un maximum d'information contenu dans l'échantillon initial.

La méthode de balancement entropique consiste à accorder une pondération à chaque observation (correspondant ici à un élève) de manière à balancer les moments (spécifiés par le chercheur) des covariables entre les groupes traités et non-traités. Les covariables en question seront sélectionnées suite à une régression logistique qui nous permettra d'isoler les déterminants de la décision de fréquenter une école privée. Le balancement nous permettra ainsi d'éliminer ou de minimiser les différences entre les groupes traités et non-traités, permettant par la suite d'isoler l'impact du type d'école fréquentée sur les résultats aux tests. Les poids ainsi générés seront récupérés pour les régressions sur chacune des cinq valeurs plausibles (qui seront discutées au chapitre traitant des données) fournies pour chacune des trois matières pour lesquelles l'élève aura été testé.

Ensuite, pour chacune des valeurs plausibles associées à chaque sujet, une régression par moindres carrés sera effectuée, avec comme variables explicatives le type d'école fréquentée ainsi qu'un vecteur de variables associées à la performance académique (niveau d'éducation des parents, classe sociale, statut d'immigration, etc.), le tout en utilisant les poids générés par le balancement entropique. La moyenne des cinq coefficients obtenus pour la variable du type d'école correspondra finalement à l'effet de l'éducation secondaire au privé sur les résultats PISA pour chacune des trois matières.

Nous effectuerons d'abord cette analyse pour l'échantillon complet. Puis, nous traiterons des sous-échantillons afin d'isoler l'effet moyen du traitement sur les traités (ATT) pour différents groupes, par exemple entre garçons et filles ou encore l'appartenance à une année scolaire spécifique.

La présentation de ce mémoire se déclinera comme suit. Nous effectuerons d'abord une revue de littérature afin de situer notre travail dans les divers courants de recherche pertinents et d'établir les bases méthodologiques de celui-ci. Nous présenterons ensuite l'origine et la nature des données utilisées, ainsi que les variables sélectionnées dans notre modèle. Nous discuterons par la suite de la méthodologie employée. Finalement, nous présenterons et interpréterons les résultats obtenus.

## CHAPITRE I

### REVUE DE LITTÉRATURE

#### 1.1 Revue des études sur l'école privée

##### 1.1.1 Impact de l'école privée sur ses élèves

Les études cherchant à estimer l'effet moyen du traitement sur les traités (désormais ATT) de l'éducation au privé sont très hétérogènes. Peu utilisent les mêmes méthodes ou les mêmes données. Celles utilisant les données PISA n'utilisent jamais plus d'une des enquêtes disponibles. Elles ne s'entendent pas sur l'existence de cet effet lorsque l'on contrôle pour différents facteurs, notamment ceux associés aux caractéristiques socio-économiques.

Par exemple, les résultats PISA de 2000 ainsi que ceux de l'Enquête auprès des jeunes en transition (EJET - cohorte A) pour le Canada entier furent utilisés dans un modèle de régression par moindres carrés ordinaires avec effets fixes de provinces, prenant en compte les caractéristiques socio-économiques des répondants ainsi que les effets de pairs (Frenette *et al.*, 2015). Une fois tous ces facteurs considérés, l'effet mesuré diminue considérablement, jusqu'à devenir nul dans la plupart des mesures (notamment les résultats obtenus aux tests PISA, mais également dans le taux de graduation secondaire et la fréquentation universitaire tels que mesurés par l'EJET).

L'étude de Frenette *et al* comporte toutefois certains problèmes lorsque nous souhaitons appliquer ses conclusions au cas qui nous intéresse. D'abord, il n'est pas surprenant de constater que l'effet fixe de provinces capture une grande partie (52,3 %) de l'écart initialement constaté entre les élèves fréquentant le privé et ceux du public. En effet, la fréquentation de l'école privée est nettement plus élevée au Québec et, à un moindre niveau, en Colombie-Britannique. Les auteurs attribuent ceci à l'application d'un curriculum provincial particulier mais n'abordent pas la question des subventions offertes aux parents au Québec qui ont pour effet de diminuer la sélection par statut socio-économique qui domine dans le reste du Canada, où les écoles privées ne sont généralement pas subventionnées (mis à part la Colombie-Britannique et, dans une moindre mesure, la Saskatchewan) et sont ainsi fréquentées par des jeunes provenant de statut socio-économiques supérieurs.

La conception de l'échantillon pourrait également être source de biais en raison de son attention portée sur la fréquentation d'institutions post-secondaires des étudiants à l'âge de 23 ans (au cours du cycle 5 de l'EJET). Le nombre de répondants chute significativement sur cette période, soulevant ainsi des doutes quant au biais de non-réponse, par exemple en affichant un taux de diplomation anormalement élevé, évalué à 95,5 % pour les élèves du public alors que le taux de diplomation moyen au Canada pour les 20-24 ans serait de 89,5 % uniquement. De plus, leur sous-échantillon des élèves du privé en 2000 est de faible taille, et le Québec compte pour environ 50 % de ce qui correspondrait à 500 élèves canadiens. C'est toutefois au Québec que le taux d'obtention d'un diplôme d'études secondaire est le plus faible, contribuant encore davantage au biais de l'échantillon (McMullen et Gilmore, 2010). Finalement, l'utilisation d'un modèle de moindres carrés ordinaires sans appariement offre souvent des résultats de piètre qualité lorsqu'on les compare à d'autres méthodes d'estimation (Vandenberghe et Robin, 2004).

De plus, un indice socio-économique souvent utilisé comme variable de contrôle

lorsque les données de PISA sont utilisées est le ESCS (*Index of economic, social and cultural status*), généré par les responsables de l'enquête. Celui-ci n'est pas sans faille et est fréquemment critiqué (Carnroy et Rothstein, 2013), entre autres en raison de son caractère arbitraire. Par exemple, l'indice accorde un poids égal aux différents éléments qui le compose, soit un indice des possessions culturelles et matérielles, un indice de classement professionnel (ISCED, décrit plus loin) ainsi que l'éducation des parents. Il semblerait que, par sa construction, cet indice soit mieux adapté aux analyses internationales qu'à celle portant sur un seul pays.

En se concentrant uniquement sur le Québec, on retrouve un effet positif et significatif de la fréquentation d'un établissement privé au secondaire pour les élèves québécois sur les scores en mathématiques obtenus aux tests CAT/2 (*Canadian Achievement Test, 2nd Edition*) tels qu'obtenus dans l'Enquête longitudinale nationale sur les enfants et les jeunes (ELNEJ) (Lefebvre *et al.*, 2011). L'échantillon de cette étude comprend des enfants entre 7 et 15 ans ayant passé au moins deux fois le test de mathématiques, une fois au primaire et l'autre au secondaire. On tente ainsi d'évaluer l'impact d'un transfert d'une école publique à une école privée, étant donné que la majorité des transferts du public au privé s'effectue à ce moment, en raison des subventions plus généreuses ainsi qu'une plus grande offre d'écoles. On observe ainsi l'effet du transfert au privé sur la progression des résultats en mathématiques entre les deux tests. Un modèle à effets fixes par enfant est utilisé afin de contrôler l'impact d'une plus grande importance accordée à l'éducation par les familles au statut socio-économiques supérieur. Le résultat obtenu démontre que la décision d'envoyer un enfant à l'école privée au moment de l'inscription au secondaire augmente son score en mathématiques de 4 à 10 percentiles et de 12 % à 36 % d'un écart-type. Ceci pourrait être expliqué par des effets de pairs ou encore par l'emploi de meilleurs enseignants par les établissements privés. Cette dernière hypothèse est toutefois discutable, étant donné

l'absence de données observables sur la qualité des enseignants.

En étudiant le cas des écoles à charte du Massachusetts (où les élèves sont assignés aux différentes écoles suite à une loterie), Angrist *et al.* (Angrist *et al.*, 2011) utilisent un modèle de régression par moindres carrés en deux étapes (*two-stage least squares*) afin d'estimer l'impact de ces écoles sur les résultats scolaires des élèves fréquentant un *middle school* (généralement l'équivalent de la sixième année du primaire à la deuxième année du secondaire dans le système québécois) ou l'école secondaire (*high school*, soit de la troisième année du secondaire à la sixième année du secondaire du système scolaire américain). Les effets sont les plus élevés en mathématiques, où les résultats des élèves des *middle school* à charte surpassent ceux des écoles traditionnelles par 25,3 % d'un écart-type, alors que l'écart pour les écoles secondaires est de 36,7 %. L'effet est moins prévalent chez les élèves plus jeunes, particulièrement en anglais où il semble nul, alors qu'il est estimé à 25,9 % à 33,4 % d'un écart-type pour les élèves du secondaire.

On cite également comme source possible de ces résultats l'aspect traditionnel ainsi que les valeurs entourant les écoles privées, qui sont souvent d'anciens collèges classiques laïcisés. Ainsi, ces écoles attireraient de meilleurs enseignants en raison du nombre moins élevé d'élèves problématiques et de la plus grande discipline qui y règnerait, de même qu'en raison de la meilleure reconnaissance de la performance des enseignants (Lefebvre *et al.*, 2011). Cette même question est également soulevée par Vandenberghe, qui remarque que l'ATT des écoles privées semble plus élevé dans les pays où les administrations de celles-ci sont soit de nature ou d'origine religieuses (Vandenberghe et Robin, 2004). Il soulève ainsi la question de « l'hypothèse religieuse », c'est-à-dire si la question de l'ATT de l'école privée ne devrait pas être celle de l'ATT de l'école religieuse.

Cette même étude offre également une conclusion intéressante dans sa comparai-

son des effets de l'école privée dans différents pays, qui sont extrêmement hétérogènes. Ceux-ci sont inévitablement négatifs en Autriche (jusqu'à un écart-type en-dessous des établissements publics, selon le sujet et la méthode d'estimation employée, ce qui est un écart considérable étant donné les effets généralement observés), mais fortement positifs au Brésil (jusqu'à 1.15 écart-type au-dessus de la moyenne du public). D'autres pays, tels que l'Espagne, le Danemark ou le Mexique, n'obtiennent aucun impact significatif, quel que soit le sujet ou la méthode d'estimation utilisée (Vandenberghe et Robin, 2004). Certains de ces résultats sont toutefois dépendants de la méthode d'estimation utilisée. Nous traiterons de ce problème plus bas.

Malgré les différences entre pays et méthodes, un constat de Vandenberghe demeure robuste : le signe des effets demeurent généralement identiques lorsqu'ils sont statistiquement significatifs, peu importe le sujet étudié ou la méthode employée (Vandenberghe et Robin, 2004). Ceci revient à dire que l'effet de l'école privée varie selon les pays, et serait influencé par le contexte qui encadre son administration, son curriculum, son historique et autres caractéristiques.

### 1.1.2 Effet d'« écrémage » et effets de pairs

La présence d'un établissement d'enseignement privé sur un territoire entraîne inévitablement une perte d'élèves pour les écoles publiques le desservant. L'hypothèse de l'effet d'écrémage suppose que cet exil touchera démesurément les élèves provenant de familles au statut socio-économique plus élevé, ainsi que de celles accordant une valeur plus grande à l'éducation de leurs enfants.

La littérature ne mène à aucun consensus sur l'impact d'un effet d'écrémage. Aux États-Unis, de nombreuses études se sont penchées sur l'impact de l'apparition d'écoles à charte (écoles non-sélectives, financées publiquement mais ayant un haut niveau de liberté aux niveaux administratifs et pédagogiques) sur les écoles

publiques de la région (Altonji *et al.*, 2010; Booker *et al.*, 2008; Imberman, 2011; Ni, 2009). Certaines études découvrent un impact positif, notamment sur les campus et les élèves en difficulté, qu'ils attribuent entre autres à l'introduction de compétition sur un marché auparavant monopolistique, ou encore à l'homogénéisation des salles de classe qui permettrait à l'enseignant d'ajuster sa technique d'enseignement au type d'élèves qui compose le groupe (Booker *et al.*, 2008). D'autres études mesurent un impact négatif (Imberman, 2011; Ni, 2009), bien que celui-ci soit généralement minime lorsque comparé aux bénéfices offerts par l'introduction de mesures favorisant l'accès aux écoles privées (Altonji *et al.*, 2010). Les questions d'ordre administratifs et de contextes social, culturel et économique se posent toutefois encore étant donné les différences régionales entre les études prouvant un effet positif (provenant du Texas) et celles trouvant un effet négatif (la plupart originaires des états du Midwest) dans le cas des écoles à charte.

Un concept voisin de celui d'effet d'écémage est celui d'effet de pairs, qui est assez facile à définir mais qui pose des défis économétriques considérables. Il s'agit de l'effet qu'ont les caractéristiques des pairs sur l'élève. On suppose qu'être entouré de collègues performants ou motivés aura un effet positif sur les résultats d'un élève donné. Les études de l'OCDE tendent également à attribuer une grande partie (près des trois quarts) de l'impact de l'éducation au privé aux effets de pairs de nature socio-économique (OCDE, 2012).

Toutefois, la littérature mesure cet effet de façon très disparate. On choisit parfois le niveau moyen d'éducation des parents pour l'école entière (Frenette *et al.*, 2015), ou encore un vecteur de covariables (socio-économiques, démographiques et habiletés) pour les étudiants d'un groupe (Boucher *et al.*, 2014). L'endogénéité de la décision de fréquentation de l'école privée rend leur implémentation délicate et controversée. Lorsque cet effet est mesuré, il est évidemment positif et capture une partie significative de l'effet de l'école privée (Frenette *et al.*, 2015) et est

particulièrement significatif en mathématiques (Boucher *et al.*, 2014).

## 1.2 Revue de méthodologie économétrique

### 1.2.1 Balancement entropique

En présence de biais de sélection qu'on suppose basé sur des caractéristiques observables, nous proposons l'utilisation de la méthode d'appariement dite de balancement entropique (Hainmueller, 2012) afin de réduire et neutraliser leur impact sur nos estimations. Nous aborderons dans le chapitre de méthodologie les détails mathématiques de cette approche.

Les méthodes d'appariement sont utilisées en recherche depuis plusieurs décennies. Par exemple, la méthode par score de propension fut initialement proposée par Rosenbaum et Rubin en 1983, et critiquées par Lalonde en 1986 (Lalonde, 1986) en raison des biais qu'elle pouvait introduire. Le processus employé est particulièrement fastidieux car il requiert du chercheur qu'il itère en la modélisation du score de propension, l'appariement et la vérification du balancement, et peut également être source de biais si le balancement des covariables n'est pas coordonné (Hainmueller, 2012).

La méthode proposée par Hainmueller est plus efficace car elle vise directement le balancement des covariables. Elle consiste à générer des ensembles de poids pour le groupe-contrôle faisant en sorte que les contraintes de balancement entre groupes traités et non-traités pour les moments et les covariables spécifiés par le chercheur soient satisfaites. L'ensemble choisi sera celui dont la distance qui le sépare de celui d'origine sera minimisée.

La méthode de balancement entropique offre certains avantages face aux autres méthodes de balancement. Son principal atout est la flexibilité de sa méthode de repondération qui permet de conserver davantage d'information que les méthodes

d'appariement par score de propension, tel que celle du plus proche voisin (*nearest neighbour matching*). En effet, on permet aux poids de varier librement plutôt que de simplement prendre la valeur 1 ou 0 selon la présence ou l'absence d'un appariement pour une observation donnée. Étant donné qu'on contraint la distance avec l'ensemble de poids d'origine, on préserve également l'efficacité de l'analyse ultérieure.

Dans le même papier, Hainmueller teste la performance de la méthode de balancement entropique avec les données de Lalonde (Lalonde, 1986). Il s'agit d'une étude classique sur les effets d'un programme de formation à l'emploi, fréquemment référencée en raison de son accès à des données expérimentales. L'ensemble de données comporte deux parties : la première provient d'une expérience randomisée sur les effets du programme, la seconde est un échantillon de la population qui serait typiquement utilisé dans une étude sur un tel programme, où on retrouve des individus faisant partie des groupes traités ainsi que non traités. Les données expérimentales permettent de chiffrer l'impact du programme, qui est ensuite comparé aux estimations sur ce même effet produites par différentes méthodes économétriques. L'application du balancement entropique sur ces données généralement jugées hautement débalancées, suivie d'un simple modèle de différence en moyennes produit un estimé très près des valeurs expérimentales et très efficace comparé aux méthodes alternatives.

### 1.2.2 Choix du modèle

Le choix du modèle utilisé dans l'analyse influence parfois le résultat obtenu. Dans leur papier, Vandenberghe et Robin obtiennent des résultats significativement différents dans le cas de certains pays selon la méthode qu'ils utilisent (Vandenberghe et Robin, 2004). En utilisant les méthodes d'appariement par score de propension (avec fonction noyau et plus proche voisin), l'effet mesuré dans le cas de la France

est faiblement positif, mais celui-ci devient fortement négatif lorsque les méthodes des variables instrumentales ou celle du modèle d'Heckman en deux étapes. Cette différence est causée par le choix économétrique de la source du biais de sélection. En effet, les méthodes par variables instrumentales ou celle d'Heckman procèdent selon l'hypothèse que le biais réside dans des caractéristiques non-observables du traité, alors que les méthodes par score de propension font l'hypothèse que les covariables utilisées capturent totalement les différences entre groupes traités et non-traités.

Il est intéressant de noter que leur utilisation des méthodes d'appariement par score de propension occasionne l'élimination de 20 % des observations du groupe traité dans certains cas (Vandenberghe et Robin, 2004), ces observations se retrouvant hors-support, c'est-à-dire sans contrepartie dans le groupe non-traité. Ceci démontre l'avantage de la méthode du balancement entropique en ce qui concerne la conservation d'informations présentes dans les données originales.

### 1.2.3 Sensibilité aux variables confondantes

Les résultats obtenus suite à l'utilisation d'une méthode d'appariement tel que celle du balancement entropique sont sujet à l'hypothèse d'indépendance conditionnelle. Il convient ainsi de tester la validité de ces résultats en invalidant cette hypothèse par le biais d'une simulation de variable confondante (Ichino *et al.*, 2008), c'est-à-dire une covariable binaire à la fois liée à la sélection du traitement ainsi qu'à la réponse au traitement. La méthode proposée par Ichino *et al.* consiste à attribuer des valeurs arbitraires mais significatives (pouvant provenir de sources empiriques) aux quatre probabilités associées à l'appartenance au groupe traité ainsi qu'à un résultat positif du traitement, en gardant constamment positive la variable binaire confondante. L'ATT est ensuite ré-estimé par plusieurs simulations, se trouvant potentiellement absorbé par la variable ainsi simulée. In-

versement, il est également possible de simuler une variable confondante tel que l'ATT devient nul, et ensuite d'évaluer la plausibilité des probabilités associées à ce scénario. Le chapitre traitant de méthodologie adressera plus en profondeur les détails économétriques de cette méthode.

## CHAPITRE II

### DONNÉES ET VARIABLES

#### 2.1 Présentation des données

Les données utilisées dans cette étude proviennent de l'enquête PISA, réalisée depuis 2000 sur une base triennale par Statistique Canada au nom de l'OCDE. L'enquête de 2000 fut réalisée conjointement avec l'Enquête auprès des jeunes en transition (EJET) et indépendamment par la suite. Dans cette première version, des questionnaires furent acheminés aux étudiants, aux parents ainsi qu'au directeur de l'école fréquentée, dont la méthode de sélection sera traitée au chapitre suivant. Le questionnaire des parents ne fut plus utilisé à partir de 2003. Les questions traitant des parents (niveau d'éducation, emploi, etc.) furent subséquemment incluses dans le questionnaire adressé aux étudiants.

Le questionnaire destiné aux étudiants aborde d'abord les questions d'ordre socio-démographique telles que l'âge, le sexe, l'année scolaire et le pays d'origine. À partir de 2003, celui-ci inclut également les questions sur les parents tel que mentionné plus haut. On y questionne également l'étudiant sur ses habitudes et comportements relatifs à l'éducation et son environnement scolaire (ambitions d'études futures, relation avec les professeurs, comportements de délinquance, etc.). Finalement, on teste ses habiletés en lecture, en mathématiques et en science. Celles-ci seront mesurées par le biais des valeurs plausibles, que nous aborderons plus bas.

Le questionnaire sur l'école est adressé à son directeur. On l'interroge sur les étudiants qui fréquentent l'établissement, les enseignants qui y sont employés (selon les matières, les qualifications et le statut d'emploi), son environnement (taille de la ville qui l'héberge) ainsi que le climat qui y règne. Parmi ces dernières questions, on retrouve les détails concernant entre autres la sélectivité de l'école, le type d'établissement (public ou privé), les mesures disciplinaires employées, le climat d'enseignement et les attentes des parents envers l'école et son personnel.

Le questionnaire des parents, tel que mentionné plus haut, ne fut employé qu'en 2000. Plusieurs informations pertinentes s'y retrouvent, tel que le revenu familial annuel et les attentes envers la performance scolaire du jeune ; informations qu'on ne pourra retrouver dans les questionnaires de 2003 et des années suivantes. Son taux de réponse était toutefois significativement plus faible que celui des autres questionnaires (environ 90 %). On utilisera tout de même les informations sur le revenu familial annuel dans notre modèle de simulation de la sensibilité de l'effet de traitement aux variables confondantes.

Il est important de noter que notre étude est soumise aux règles de divulgation et de confidentialité du Centre interuniversitaire québécois de statistique sociale et de Statistique Canada. Ainsi, certaines statistiques ne peuvent être dévoilées, telles que les minima et maxima. De plus, des poids d'échantillonnage (que nous traiterons au chapitre de méthodologie) fournis par Statistique Canada sont utilisés tout au long de l'analyse afin de dresser un portrait fidèle de la population générale visée tout en prenant en compte la non-réponse de certains participants. Le nombre d'observations est donc toujours fourni à des fins d'information, mais seules les fréquences ont une signification statistique.

## 2.2 Valeurs plausibles

Les valeurs plausibles sont générées par les responsables de l'enquête PISA sur la base des résultats obtenus par les répondants et proviennent de distributions se chevauchant générées préalablement. Une fois le score de l'étudiant obtenu, cinq valeurs plausibles provenant des distributions auxquelles ce score appartient sont tirées. Ce processus cherche à minimiser l'impact des conditions mentales et physiques de l'étudiant au moment de passer le test, entre autres facteurs pouvant affecter sa performance. Néanmoins, les responsables de l'enquête affirment que les valeurs plausibles sont, dans l'agrégé, représentatives de la population étudiée.

## 2.3 Cohérence des données

Étant donné la nature cyclique de l'enquête PISA, nous avons dû effectuer certains choix afin de préserver la cohérence des données à travers les échantillons dans le temps. À cet égard, l'enquête de 2000 mérite une considération particulière. L'utilisation du questionnaire des parents fit en sorte que les informations relatives à leurs niveaux d'éducation furent catégorisées différemment de la classification standardisée ISCED (*International Standard Classification of Education*) utilisée dans les années subséquentes. Afin de palier à cette déficience, nous utiliserons une variable pour l'éducation parentale en 2000 et une autre post-2000, utilisant des catégorisations différentes quant aux études post-secondaires.

Il importe de noter que les taux de réponse au Québec furent inférieurs aux taux normalement requis par les concepteurs de l'enquête PISA pour les enquêtes de 2009 et 2012. Toutefois, après inspection des données, la qualité fut jugée suffisante pour que les résultats puissent être utilisés. Il semblerait également que les familles de milieux défavorisés aient refusé que leurs enfants participent dans une plus grande proportion que celles de milieux plus aisés. Ceci fait en sorte que

s'il y a biais dans nos estimations, il mènera à un estimé plus conservateur de l'effet de traitement car les scores moyens des élèves du publics seront plus élevés, étant donné que les élèves de milieux défavorisés sont sur-représentés dans ces établissements.

À chaque année, l'enquête PISA se concentre sur une des trois matières qu'elle évalue. De 2003 à 2012, cela signifie un test plus approfondi sur la matière-cible de cette enquête, et un test plus restreint pour chacune des deux autres matières. Toutefois, l'approche fut différente en 2000, où on choisit de faire passer à tous les étudiants un test de lecture, alors que moins d'étudiants subirent les tests de mathématiques et de science.

D'année en année, les questionnaires abordent les sujets de manière quelque peu différente, menant parfois à des problèmes de codification des variables. Par exemple, lorsqu'on demande à un étudiant d'estimer le nombre de livres qui se trouvent chez lui, le questionnaire utilisé en 2000 ne fournissait pas les mêmes intervalles que celui utilisé dans les années subséquentes. Plusieurs variables ont dû être ré-encodées afin de maintenir la cohérence des données à travers les différentes années.

Comme les questionnaires sont remplis de manière non-coercitive, certains répondants ne répondent pas à toutes les questions. Afin de maintenir une taille échantillonnaire acceptable, plutôt que de ne pas compter ces observations, nous avons ré-encodé ces réponses manquantes par la valeur « 0 » lorsque ceci était applicable. Dans certains cas (par exemple dans le cas des quintiles de statut socio-économique, dont nous discuterons plus bas), nous avons éliminé complètement les observations manquantes de notre analyse.

Afin de permettre à notre modèle de balancement entropique de converger, nous avons dû éliminer certaines observations aberrantes. Par exemple, la quasi-totalité

des élèves se trouvant dans une année scolaire inférieure à la troisième année du secondaire se retrouvaient dans des établissements publics; et inversement, les quelques élèves se trouvant en cinquième année du secondaire fréquentaient pratiquement tous un établissement privé. Conserver ces observations nous faisait risquer l'introduction de biais en raison des poids anormalement élevés qui auraient dû être assignés aux rares observations appartenant au groupe opposé; ou empêchait l'algorithme de balancement entropique de converger. Toutes les statistiques descriptives présentées dans ce mémoire proviennent de l'échantillon ainsi restreint.

## 2.4 Variables et statistiques descriptives

### Type d'établissement

En consultant le tableau 2.1, on remarque une hausse de la fréquentation de l'école privée au Québec en 2009 et en 2012, passant de 18,42 % en 2000 à 23,94 % en 2012. En moyenne, 20,18 % des élèves de notre échantillon fréquentent un établissement d'éducation privé.

**Tableau 2.1** Type d'établissement par année

	Public		Privé		Total	
	n	%	n	%	n	%
2000	2612	81,58	590	18,42	3202	18,77
2003	2606	82,97	535	17,03	3141	18,41
2006	2888	80,88	683	19,12	3571	20,94
2009	2878	78,10	807	21,90	3685	21,61
2012	2629	76,08	828	23,94	3457	20,27
Total	13614	79,82	3442	20,18	17056	100

Source : Calcul de l'auteur à partir des données pondérées de PISA

## Scores PISA

Le tableau 2.2 démontre l'écart de performance entre les élèves fréquentant un établissement privé et ceux qui évoluent dans le système public. On remarque *a priori* que la différence est plus marquée pour les mathématiques, et l'est moins pour la science. L'écart-type des résultats est toujours plus élevé pour les écoles publiques que pour les écoles privées, vraisemblablement en raison de l'universalité du système.

**Tableau 2.2** Scores PISA : public et privé

	Moyenne	Écart-type
<b>Mathématiques</b>		
Tous	553,90	78,29
Public	544,05	77,39
Privé	592,97	69,14
<b>Lecture</b>		
Tous	538,07	81,05
Public	528,61	80,76
Privé	575,51	70,69
<b>Science</b>		
Tous	538,66	81,70
Public	529,74	81,41
Privé	574,03	72,77

Source : Calcul de l'auteur à partir des données pondérées de PISA

Le tableau 2.3 illustre les écarts de performance aux différents tests PISA entre les filles et garçons. L'écart le plus marqué se trouve en lecture, où les filles obtiennent en moyenne 27,83 points de plus que les garçons, et est moins marqué en science où ces derniers ont 9,45 points d'avance. Les écarts sont toutefois peu significatifs,

se mesurant au plus à 34,34% d'un écart-type au plus dans le cas des scores en lecture.

**Tableau 2.3** Scores PISA : filles et garçons

	Moyenne	Écart-type
<b>Mathématiques</b>		
Tous	553,90	78,29
Filles	548,38	75,98
Garçons	561,90	79,91
<b>Lecture</b>		
Tous	538,07	81,05
Filles	551,54	76,38
Garçons	523,71	83,39
<b>Science</b>		
Tous	538,66	81,70
Filles	533,93	79,44
Garçons	543,68	83,73

Source : Calcul de l'auteur à partir des données pondérées de PISA

### Sexe de l'étudiant

En consultant le tableau A.4 en Annexe A, on constate que la fréquentation de l'école privée est plus élevée chez les filles que chez les garçons.

### Taille de la communauté

Le tableau A.5 en Annexe A démontre l'écart de fréquentation d'une école privée selon la taille de communauté où celle-ci se retrouve. Près de la moitié des élèves étudiant au privé fréquentent une école située dans une région urbaine (plus de

100 000 habitants), alors qu'uniquement 33,71 % des élèves au total fréquentent une école dans ce type de communauté.

#### Statut d'immigration des parents

L'origine des parents semble un facteur important concernant la décision pour l'enfant de fréquenter un établissement privé selon le tableau A.6 en Annexe A. La fréquentation d'un établissement privé passe de 19,05 % jusqu'à 25,65 % lorsqu'au moins un des deux parents est né à l'extérieur du Canada. Ceci tend à donner du poids aux arguments selon lesquels l'impact des familles sur les résultats scolaires réside dans des facteurs autres que simplement économiques.

#### Statut socio-économique

En catégorisant par quintile le plus élevé l'indices ISEI des parents (*International Socio-Economic Index of Occupational Status*, qui classe les professions selon le niveau d'éducation requis et le revenu moyen de la profession) (Ganzeboom et Treiman, 1996), nous obtenons une variable de statut socio-économique (SSE). Nous avons sélectionné cette variable comme étant le substitut le plus adéquat au revenu de la famille, qui n'était plus disponible après 2000, mais qui est aussi un indice valorisé dans la littérature du statut socio-économique. Le tableau 2.4 illustre la fréquentation de chaque type d'établissement selon le SSE des parents. On remarque un lien de corrélation marqué entre le statut professionnel le plus élevé des parents et la fréquentation d'un établissement privé.

#### Éducation de la mère

Nous avons choisi l'éducation de la mère comme indicateur d'éducation des parents pour des raisons socio-démographiques. En effet, dans la majorité des cas, le « chef de famille » dans le cas d'une famille monoparentale se trouve à être la mère,

**Tableau 2.4** Quintile de statut socio-économique selon le type d'établissement

	1er quintile		2e quintile		3e quintile	
	n	%	n	%	n	%
Public	3174	92,99	2961	86,83	2852	83,58
Privé	239	7,01	449	13,17	560	16,42
	4e quintile		5e quintile			
	n	%	n	%		
Public	2508	73,56	2119	62,11		
Privé	902	26,44	1292	37,89		
Observations	17056					

Source : Calcul de l'auteur à partir des données pondérées de PISA

qui obtient généralement la garde de l'enfant. Nous jugeons donc que le niveau d'éducation de la mère se trouve à être le meilleur indicateur du niveau d'éducation des parents.

Tel que mentionné plus haut, les données concernant l'éducation parentale sont codées différemment en 2000 des autres années. Nous avons donc généré deux variables selon l'enquête à laquelle l'observation appartient. Les tableaux 2.5 et 2.6 illustrent le type d'établissement fréquenté par l'élève selon le niveau d'éducation de la mère. On constate que l'obtention d'un diplôme post-secondaire est un facteur déterminant dans le choix du type d'établissement fréquenté.

#### Nombre de livres à la maison

Nous avons choisi le nombre de livres à la maison comme indicateur de possessions culturelles domestiques. Nous posons l'hypothèse que les familles offrant davantage de matériel culturel aux enfants accordent une valeur plus grande à leur développement intellectuel, favorisant leur succès scolaire par ricochet. Il s'agit

**Tableau 2.5** Éducation de la mère en 2000 selon le type d'établissement

		n	% de la sous-catégorie	% total
Données manquantes	Public	591	83,71	16,17
	Privé	115	16,29	3,15
	Total	706	100	19,32
Secondaire non-complété	Public	416	93,19	11,37
	Privé	30	6,81	0,83
	Total	446	100	12,20
Secondaire complété	Public	897	86,50	24,55
	Privé	140	13,50	3,83
	Total	1037	100	28,38
Diplôme technique / professionnel / collégial	Public	679	80,35	18,58
	Privé	166	19,65	4,55
	Total	845	100	23,13
Universitaire premier cycle	Public	331	67,33	9,07
	Privé	161	32,67	4,40
	Total	492	100	13,47
Universitaire cycles supérieurs	Public	67	52,53	1,84
	Privé	61	47,47	1,66
	Total	128	100	3,50
Observations		3655		

Source : Calcul de l'auteur à partir des données pondérées de PISA

**Tableau 2.6** Éducation de la mère post-2000 selon le type d'établissement

		n	% de la sous-catégorie	% total
Données manquantes	Public	203	91,24	1,51
	Privé	19	8,76	0,15
	Total	222	100	1,66
Secondaire non-complété	Public	1006	93,85	7,51
	Privé	66	6,15	0,49
	Total	1072	100	8,00
Secondaire complété	Public	1918	88,90	14,31
	Privé	239	11,10	1,79
	Total	2157	100	16,10
Diplôme professionnel	Public	2103	87,95	15,69
	Privé	288	12,05	2,15
	Total	2391	100	17,84
Diplôme collégial	Public	2897	79,62	21,62
	Privé	741	20,38	5,53
	Total	3638	100	27,15
Diplôme universitaire	Public	2515	64,16	18,77
	Privé	1405	35,84	10,48
	Total	3920	100	29,25
Observations		13401		

Source : Calcul de l'auteur à partir des données pondérées de PISA

également d'un complément à notre indicateur de classe sociale, traduisant un niveau de richesse familiale plus élevé.

#### Langue parlée à la maison

En observant le tableau A.7 en Annexe A, on remarque un taux de fréquentation d'école privée étonnamment bas pour les élèves parlant une langue autre que le français ou l'anglais à la maison, ce qui est d'autant plus choquant que nous avons démontré que les familles immigrantes fréquentaient davantage ce type d'établissement. Il y aurait donc un clivage entre immigrants francophones et anglophones d'une part, et ceux allophones de l'autre.

#### Âge et année scolaire

En raison des facteurs touchant le redoublement ainsi que les naissances tardives (la règle étant que les enfants nés après le 30 septembre soient considérés comme faisant partie de la cohorte de l'année suivante), nous incluons l'âge et l'année scolaire de référence de l'étudiant comme variable (voir tableaux A.1, A.2 et A.3 en Annexe A).

## CHAPITRE III

### MÉTHODOLOGIE

#### 3.1 Méthode échantillonnale

Afin de dresser un portrait adéquat de la population étudiée, l'enquête PISA procède en deux étapes pour son échantillonnage : la sélection des écoles participantes, puis le choix des élèves questionnés.

La sélection des écoles se fait selon des probabilités proportionnelles à leur taille (PPT). Ainsi, les écoles plus fréquentées ont davantage de chance d'être sélectionnées. Sachant que 35 étudiants par école sélectionnée doivent répondre au questionnaire ainsi que la taille de l'échantillon souhaité, le nombre d'écoles à sélectionner est déterminé. On calcule ensuite un intervalle échantillonnal correspondant au ratio du nombre d'étudiants dans la population sur le nombre d'écoles dans l'échantillon final. Puis, un nombre aléatoire d'une distribution  $[0;1]$  est tiré et multiplié par l'intervalle échantillonnal. Ayant au préalable trié les étudiants par école, ce nombre correspondra à l'un d'entre eux. L'école qu'il fréquente sera la première école de l'échantillon. L'intervalle échantillonnal sera additionné au nombre obtenu précédemment, et si l'élève correspondant à celui-ci appartient à une autre école, celle-ci est ajoutée à l'échantillon. Ce processus est répété jusqu'à l'obtention du nombre d'écoles requis (OCDE, 2009).

Tel que mentionné précédemment, 35 étudiants par école pré-sélectionnée doivent

être soumis au questionnaire PISA. La probabilité pour ceux-ci d'être choisis est donc inversement proportionnelle à la taille de l'école. Le tableau 3.1 donne un exemple des probabilités associées à un scénario où on compte une population de 400 étudiants dans 10 écoles, parmi lesquelles 4 seront sélectionnées et 10 élèves par école seront interrogés.

**Tableau 3.1** Probabilités de sélection d'école, intra-école et finale et poids associés pour un échantillon PPT avec écoles de tailles inégales

École	Taille de l'école	Probabilité de l'école	Poids de l'école	Probabilité intra-école	Poids intra-école	Probabilité finale de l'élève	Poids final de l'élève	Somme des poids finaux
1	10	-	-	-	-	-	-	-
2	15	-	-	-	-	-	-	-
3	20	0,2	5,00	0,500	2,0	0,1	10	100
4	25	-	-	-	-	-	-	-
5	30	-	-	-	-	-	-	-
6	35	-	-	-	-	-	-	-
7	40	0,4	2,50	0,250	4,0	0,1	10	100
8	45	-	-	-	-	-	-	-
9	80	0,8	1,25	0,125	8,0	0,1	10	100
10	100	1,0	1,00	0,100	10	0,1	10	100
Total	400		9,75					400

## 3.2 Poids d'échantillonnage et poids répliqués

### 3.2.1 Poids d'échantillonnage

Les poids d'échantillonnage utilisés par l'enquête PISA ont pour but de représenter adéquatement les différents segments de la population étudiée. Ceux-ci sont inversement proportionnels à la probabilité d'être sélectionné, à la fois pour l'école dans la population ainsi que pour l'élève à l'intérieur de l'école. Par conception, la somme des poids des élèves est égale au nombre d'élèves dans la population, alors que la somme des poids des écoles n'est pas soumise à cette contrainte (OCDE, 2009) Le tableau 3.1 illustre les poids des écoles et des élèves dans le cas décrit plus haut.

Dans ce scénario, les poids des élèves demeurent constants. Toutefois, plusieurs circonstances font en sorte que ce n'est pas le cas en pratique. Ainsi, si certains

élèves dans une école sont absents ou refusent d'être testés, les poids des autres élèves seront proportionnellement augmentés de telle manière que la somme finale des poids demeure constante. La même logique est appliquée lorsque des écoles refusent de participer à l'enquête (OCDE, 2009).

Les poids peuvent également varier selon les intentions des concepteurs de PISA. Par exemple, si 10 % des élèves proviennent d'un milieu rural et 90 % d'un milieu urbain et que l'on souhaite comparer les résultats de ces deux groupes, il est possible que les responsables décident de sur-représenter les élèves du milieu rural afin de s'assurer de la qualité de l'information obtenue. Les poids devront évidemment être ajustés en conséquence, de telle sorte que l'échantillon représente adéquatement la distribution des élèves dans la population selon leur milieu (OCDE, 2009).

### 3.2.2 Poids répliqués

La méthode d'échantillonnage décrite précédemment, bien qu'efficace, ne permet pas de calculer l'écart-type de la même façon que nous le ferions pour un échantillon tiré au sort. En effet, nous ne pouvons considérer chacun des élèves comme une observation indépendante car ceux-ci partagent les caractéristiques communes des écoles qu'ils fréquentent (type d'établissement, taille de la communauté, statut socio-économique dans une certaine mesure, *etc.*). L'incertitude quant aux statistiques que nous calculons est donc nécessairement plus grande. La méthode des poids par répliques équilibrées répétées (*balanced repeated replication*, ou BRR) avec ajustement de Fay est donc employée, tel que suggéré par les concepteurs de PISA, afin de calculer adéquatement les écarts-types de notre échantillon.

Les poids BRR sont générés en stratifiant l'échantillon en paires d'unités d'échantillonnage primaire (*primary sampling units*, ou PSU), dans ce cas-ci les écoles qui seront réparties en 80 strates. Les estimés de ces strates sont ensuite conduits en

accordant aléatoirement à une des deux écoles un poids de  $k$  compris entre 0 et 1, et à l'autre un poids de 2 moins  $k$ . On somme ensuite les carrés de l'écart entre l'estimé de la strate et celui de l'échantillon pour les  $G$  strates, que l'on pondère ensuite par  $\frac{1}{G}$  (voir l'équation 3.1).

$$\sigma_{(\hat{\theta})}^2 = \frac{1}{G(1-k)^2} \sum_{i=1}^G (\hat{\theta}_{(i)} - \hat{\theta})^2 \quad (3.1)$$

où  $G$  est le nombre de strates, soit 80,  $k$  est le facteur de déflation de Fay, soit 0.5,  $\hat{\theta}_{(i)}$  est l'estimé répliqué et  $\hat{\theta}$  est celui de l'échantillon. Les poids sont ensuite générés par les concepteurs de l'enquête de façon à refléter cette variance en accordant les poids correspondants à chaque observation.

### 3.3 Régression logistique

La régression logistique est un modèle non-linéaire de choix discret, c'est-à-dire qu'il s'agit d'un type de régression qui modélise le choix d'un agent parmi un ensemble fini d'alternatives. Dans le cas qui nous intéresse, nous cherchons à déterminer quels sont les facteurs menant au choix de fréquenter une école privée. Cette méthode nous permettra d'identifier quelles variables sont les déterminants susceptibles d'influencer cette décision.

L'équation 3.2 illustre la probabilité que  $Y=1$  étant donné un vecteur de variables  $X$  selon la fonction de distribution cumulative logistique standard, dénotée  $F$ . Les coefficients sont obtenus par la méthode du maximum de vraisemblance.

$$\begin{aligned} \Pr(Y = 1 | X_1, X_2, \dots, X_k) &= F(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k) \\ &= \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}} \end{aligned} \quad (3.2)$$

### 3.4 Balancement entropique

La méthode du balancement entropique apporte une solution aux différences entre groupes traités et non-traités qui résultent des différents biais de sélection. Elle permet de spécifier un vecteur de covariables et de balancement les moments de celles-ci selon les choix du chercheur. L'algorithme générera ensuite les ensembles de poids permettant de converger vers une différence asymptotique de 0 entre les deux groupes et conservera l'ensemble dont la distance qui le sépare de celui d'origine est minimisée.

L'équation 3.3 représente la fonction de perte, c'est-à-dire la perte d'information occasionnée par la modification des poids d'origine de l'échantillon. La fonction  $h(\omega_i)$  est donc une mesure de distance, où  $\omega_i$  représente chaque poids individuel.

$$\min_{\omega_i} H(\omega) = \sum_{\{i|D=0\}} h(\omega_i) \quad (3.3)$$

Cette équation est sujette aux conditions de balance et de normalisation des équations 3.4, 3.5 et 3.6.

L'équation 3.4 décrit les contraintes de balancement sur les moments des covariables suivant les spécifications du chercheur.

$$\sum_{\{i|D=0\}} \omega_i c_{ri}(X_i) = m_r \quad \text{où } r \in 1, \dots, R \quad (3.4)$$

Les équations 3.5 et 3.6 décrivent les conditions de normalisation des poids.

$$\sum_{\{i|D=0\}} \omega_i = 1 \quad (3.5)$$

$$\omega_i \geq 0 \quad \forall i \quad \text{tel que } D = 0 \quad (3.6)$$

Le problème d'optimisation nous donne le Lagrangien de l'équation 3.7, où  $Z$  est un vecteur de multiplicateurs de Lagrange pour les conditions de balancement, et  $\lambda_0 - 1$  celui des contraintes de normalisation. La condition de premier ordre  $\frac{\partial L^p}{\partial \omega_i}$  nous donne la solution en 3.8.

$$\begin{aligned} \min_{W, \lambda_0, Z} L^p = & \sum_{\{i|D=0\}} \omega_i \log (\omega_i/q_i) + \sum_{r=1}^R \lambda_r \left( \sum_{\{i|D=0\}} \omega_i c_{ri}(X_i) - m_r \right) \\ & + (\lambda_0 - 1) \left( \sum_{\{i|D=0\}} \omega_i - 1 \right) \end{aligned} \quad (3.7)$$

$$\omega_i^* = \frac{q_i \exp(-\sum_{r=1}^R \lambda_r c_{ri}(X_i))}{\sum_{\{i|D=0\}} q_i \exp(-\sum_{r=1}^R \lambda_r c_{ri}(X_i))} \quad (3.8)$$

En ramenant 3.8 dans notre Lagrangien, nous obtenons l'équation 3.9, qui est non-contrainte. Un algorithme de type Levenberg-Marquardt est ensuite utilisé afin de trouver  $Z^*$  pour ce nouveau problème. Il s'agit d'une approche itérative employant une variante de la méthode de Gauss-Newton pour trouver la solution d'un problème de minimisation.

$$\min_Z L^d = \log \left( \sum_{\{i|D=0\}} q_i \exp \left( -\sum_{r=1}^R \lambda_r c_{ri}(X_i) \right) \right) + \sum_{r=1}^R \lambda_r m_r \quad (3.9)$$

### 3.5 Régressions par moindres carrés

La structure de l'enquête PISA nous contraint à utiliser la méthode de régression par moindres carrés. Une fois les poids obtenus par le balancement entropique,

ceux-ci sont utilisés lorsque chacune des cinq valeurs plausibles générées pour chaque matière est régressée sur le vecteur des variables explicatives, et l'écart-type est calculé à l'aide des poids répliqués décrits plus haut. Étant donné que les poids répliqués et les poids d'échantillonnage (sur lesquels sont basés les poids proportionnels utilisés par la suite) fournis par les responsables de l'enquête PISA et Statistique Canada sont conçus afin d'éliminer les problèmes d'hétéroscédasticité et d'autocorrélation résultant de la méthode d'échantillonnage en deux étapes, nous pouvons utiliser les moindres carrés ordinaires (MCO).

La méthode du MCO, tel qu'illustré par le modèle de l'équation 3.10, consiste à minimiser la somme des carrés des termes d'erreur. Le vecteur des estimateurs est donc celui qui minimise le carré de la différence entre la variable dépendante observée et celle que le modèle prédit (voir l'équation 3.11)

$$Y = X\beta + \epsilon \quad (3.10)$$

$$\hat{\beta} = \operatorname{argmin} (Y - X\beta)' (Y - X\beta) \quad (3.11)$$

Sous les hypothèses d'homoscédasticité (équation 3.12) et d'absence d'autocorrélation (équation 3.13).

$$E [\epsilon_i^2 | X] = \sigma^2 \quad (3.12)$$

$$E [\epsilon_i \epsilon_j | X] = 0 \quad \forall i \neq j \quad (3.13)$$

La solution du problème de minimisation nous donne notre vecteur d'estimateurs tel que défini par l'équation 3.14.

$$\hat{\beta} = (X'X)^{-1} X'Y \quad (3.14)$$

### 3.6 Effets de pairs

En raison de la conception de l'enquête PISA et des données qui sont disponibles, il a été décidé de ne pas considérer les effets de pairs dans le modèle. L'enquête ne fournit pas d'information concernant l'appartenance d'un élève à une classe ou groupe plutôt qu'un autre à l'intérieur d'une école. De plus, dans le cadre des données utilisées ici, les élèves ont 15 ans et fréquentent la même école depuis trois ou quatre ans en général. La théorie derrière l'utilisation des effets de pairs comme variable explicative dans un modèle comme le nôtre repose sur la notion que les camarades de classe d'un élève influenceront son apprentissage. Sans information sur les classes et après trois ou quatre ans de fréquentation d'une école, il est difficile d'isoler un effet de pairs. Il apparaît inapproprié et injustifié d'utiliser un substitut tel que le score moyen de l'école ou le niveau moyen d'éducation des parents afin de capter cet effet. Les effets de pairs peuvent également être considéré comme étant un facteur médiateur, c'est-à-dire qu'ayant un lien causal avec la variable de traitement (la fréquentation de l'école privée), il serait inapproprié d'inclure cet effet dans notre modèle.

### 3.7 Simulation de variable confondante

La simulation de variable confondante est une méthode utilisée afin de valider la mesure de l'ATT. Elle suppose l'existence d'une covariable binaire  $U$  non observée à la fois corrélée avec la sélection pour le traitement  $T$  et son effet  $Y$ . On suppose ainsi que l'assignation au traitement ne dépend pas uniquement des covariables observables  $W$  (équation 3.15), mais que l'hypothèse d'indépendance conditionnelle est respectée lorsqu'on prend en compte une covariable non-observable  $U$

(équation 3.16).

$$\Pr(T = 1|Y_0, Y_1, W) \neq \Pr(T = 1|W) \quad (3.15)$$

$$\Pr(T = 1|Y_0, Y_1, W, U) = \Pr(T = 1|W, U) \quad (3.16)$$

La méthode consiste à assigner des probabilités  $p_{i,j}$  lorsque  $U = 1$  décrivant chacun des quatre groupes définis par leur statut de traité ou non-traité ainsi que le résultat (*outcome*) de l'individu, traité ou non (équation 3.17).

$$\Pr(U = 1|T = i, Y = j, W) = \Pr(U = 1|T = i, Y = j) \equiv p_{ij} \quad \text{où } i, j \in \{0, 1\} \quad (3.17)$$

L'assignation des probabilités est laissé à la discrétion du chercheur et peuvent provenir d'autres sources ou être composées de façon à évaluer les effets de certains scénarios jugés plausibles sur l'ATT. Chaque individu se voit enfin attribuer une valeur de  $U$  selon son appartenance aux quatre groupes définis plus tôt et l'ATT est ré-estimé par la méthode choisie par le chercheur. Le processus est répété un nombre de fois arbitraire (ici, cent fois) et l'ATT moyen est calculé et peut être évalué.

Deux effets sont également calculés par cette méthode grâce à des régressions logistiques effectuées à chaque itération. Le premier,  $\Gamma$ , est l'effet de résultat et est obtenu suite à l'estimation du modèle  $\Pr(Y = 1|T = 0, U, W)$ .  $\Gamma$  représente l'impact de la variable  $U$  sur la probabilité d'obtenir un résultat positif (équation 3.18). Le deuxième effet est celui de sélection, noté  $\Lambda$  et obtenu suite à l'estimation de  $\Pr(T = 1|U, W)$ , et qui représente l'effet de  $U$  sur la probabilité de choisir

le traitement (3.19). Les estimations de  $\Gamma$  et  $\Lambda$  caractérisent  $U$  et assistent le chercheur dans son interprétation de l'ATT tel qu'obtenu par la simulation (Ichino *et al.*, 2008).

$$\Gamma \equiv \frac{\frac{\Pr(Y=1|T=0,U=1,W)}{\Pr(Y=0|T=0,U=1,W)}}{\frac{\Pr(Y=1|T=0,U=0,W)}{\Pr(Y=0|T=0,U=0,W)}} \quad (3.18)$$

$$\Lambda \equiv \frac{\frac{\Pr(T=1|U=1,W)}{\Pr(T=0|U=1,W)}}{\frac{\Pr(T=1|U=0,W)}{\Pr(T=0|U=0,W)}} \quad (3.19)$$

Dans notre modèle, nous utiliserons les statistiques du revenu familial comme variable confondante, qui sont disponibles exclusivement pour l'enquête réalisée en 2000 dans le cadre de l'EJET. Nous aurons donc  $U = 1$  lorsque le revenu familial sera au-dessus du revenu médian d'abord, et ensuite lorsque le revenu familial sera dans le quatrième quartile. Le résultats sera mesuré de la même manière, c'est-à-dire qu'on aura  $Y = 1$  d'abord lorsque le score dans la matière étudiée sera au-dessus du score médian, puis ensuite lorsqu'il fera partie du quatrième quartile.

## CHAPITRE IV

### PRÉSENTATION DES RÉSULTATS

#### 4.1 Estimations

##### 4.1.1 Régression logistique

Le tableau 4.1 présente les résultats de la régression logistique décrivant la probabilité de fréquentation d'une école privée. Les coefficients sont à interpréter comme une modification de probabilité par rapport au coefficient de la variable de base omise. Par exemple, le premier quintile de statut socio-économique est omis, et les coefficients des quintiles suivants représente l'augmentation de la probabilité de fréquentation d'une école privée par rapport au premier quintile. Tel que mentionné dans le chapitre des données, les informations sur les diplômes obtenus font référence à l'éducation de la mère.

Nous constatons ainsi que le statut socio-économique est hautement prédictif de la probabilité de fréquenter une école privée. Il en est de même pour l'éducation post-secondaire de la mère. Le nombre de livres à la maison est aussi particulièrement significatif une fois le seuil de 100 livres dépassé, ce qui donne du crédit à l'hypothèse mentionne au chapitre 2, c'est-à-dire que les familles offrant davantage de matériel culturel aux enfants accorderait une plus grande valeur à leur éducation. L'effet de la ville est aussi significatif, ce qui n'est pas surprenant étant donné qu'il s'agit de l'environnement où se retrouve la majorité des établissements

Tableau 4.1 Régression logistique - probabilité de fréquenter une école privée

Variable	Coefficient	Écart-type	Statistique z
2003	-0,48	0,3	-1,57
2006	-0,48	0,3	-1,58
2009	-0,33	0,3	-1,1
2012	-0,38	0,3	-1,25
Garçon	-0,04	0,05	-0,89
Ville	0,72***	0,05	13,91
Un parent immigrant	0,18**	0,08	2,24
Deux parents immigrants	0,41***	0,08	5,33
2e quintile SSE	0,48***	0,1	4,77
3e quintile SSE	0,68***	0,1	6,9
4e quintile SSE	1,01***	0,1	10,57
5e quintile SSE	1,39***	0,1	14,53
2000, secondaire non complété	-0,64**	0,24	-2,74
2000, secondaire complété	-0,19	0,15	-1,23
2000, diplôme technique/professionnel/collégial	0,12	0,15	0,79
2000, universitaire 1er cycle	0,4***	0,16	2,55
2000, universitaire cycles supérieurs	0,7***	0,22	3,14
Post-2000, secondaire non complété	-0,23	0,31	-0,72
Post-2000, secondaire complété	0,09	0,29	0,3
Post-2000, diplôme professionnel	0,2	0,29	0,69
Post-2000, diplôme collégial	0,55**	0,28	1,97
Post-2000, diplôme universitaire	0,96***	0,28	3,41
1-10 livres à la maison	0,01	0,22	0,03
11-100 livres à la maison	0,35*	0,2	1,73
101-500 livres à la maison	0,73***	0,2	3,59
>500 livres à la maison	0,8***	0,21	3,76
Anglais parlé à la maison	-0,59***	0,06	-9,86
Autre langue parlée à la maison	-0,79***	0,11	-6,92
Élève d'au moins 16 ans	-0,2***	0,05	-4,14
Élève de secondaire 4	0,47***	0,06	8,33
Constante	-3,22***	0,24	-13,28

n=17056, p-value : \*<0,10; \*\*<0,05; \*\*\*<0,01

privés. Les familles parlant une autre langue que le français à la maison a un effet négatif significatif, ce qui pourrait potentiellement être lié à une offre plus faible d'établissements privés en anglais lié à la baisse de leur poids démographique. Finalement, l'âge et l'année scolaire ont tous deux des effets relativement intuitifs.

#### 4.1.2 Estimations générales

Le tableau 4.2 résume l'effet de traitement estimé pour chaque matière. On constate ainsi un effet positif de l'éducation au privé, allant jusqu'à 26,70 points dans le cas des mathématiques. L'effet le moins marqué se trouve en science avec 23,71 points, alors que les scores lecture bénéficie d'un effet de 25,61 points. Tous les coefficients estimés sont significatifs à 99%. En terme de pourcentage d'écart-type, l'effet de traitement en mathématique est de 34,1% d'un écart-type, 31,6% en lecture et 29,02% en science.

**Tableau 4.2** Effet de traitement par matière

Matière	Coefficient	Écart-type	Statistique t	n	R <sup>2</sup>
Mathématiques	26,70***	3,10	8,61	15412	0,24
Lecture	25,61***	3,87	6,63	17056	0,22
Science	23,71***	3,49	6,80	15418	0,22
p-value : * < 0,10 ; ** < 0,05 ; *** < 0,01					

Les tableaux A.8, A.9 et A.10 en annexe détaillent les résultats de nos régressions.

#### 4.1.3 Estimations par sexe

Les tableaux 4.3 et 4.4 résument les estimations effectuées par sexe. L'effet le plus marqué se trouve en mathématiques pour les filles, qui bénéficient d'un effet de traitement équivalent à 28,48 points, et le moins marqué se trouve en science pour les garçons avec 21,93 points. En terme d'écart-type, le gain pour les filles se

mesure à 37,48% en mathématiques, 34,75% en lecture et 31,53% en science. Pour les garçons, il s'élève à 31,65% en mathématiques, 30,25% en lecture et 26,84% en science.

**Tableau 4.3** Effet de traitement par matière : filles

Matière	Coefficient	Écart-type	Statistique t	n	R <sup>2</sup>
Mathématiques	28,48***	3,75	7,60	8130	0,24
Lecture	26,54***	4,19	6,34	8971	0,21
Science	25,05***	3,60	6,95	8130	0,23
p-value : * < 0,10 ; ** < 0,05 ; *** < 0,01					

**Tableau 4.4** Effet de traitement par matière : garçons

Matière	Coefficient	Écart-type	Statistique t	n	R <sup>2</sup>
Mathématiques	24,78***	3,94	6,28	7282	0,23
Lecture	24,52***	5,35	4,59	8085	0,19
Science	21,93***	4,83	4,54	7288	0,21
p-value : * < 0,10 ; ** < 0,05 ; *** < 0,01					

#### 4.2 Simulation de variable confondante

Les tableaux 4.5, 4.6 et 4.7 illustrent les résultats obtenus par la méthode simulation de variable confondante. Tel que mentionné au chapitre de méthodologie, les statistiques du revenu familial disponibles dans l'enquête de 2000 ont été utilisées comme variable confondante dans le modèle. Dans les tableaux suivants, la colonne intitulée statistique de sélection correspond à la mesure employée comme seuil de sélection au traitement (par exemple un revenu familial au-dessus du revenu médian), et la colonne titrée statistique de résultat est celle utilisée pour l'appartenance au groupe ayant subi une issue (ou résultat, *outcome* en anglais)

positif, par exemple un score appartenant au 4e quartile. L'effet de sélection mesure l'effet de la variable confondante  $U$  sur la probabilité du choix du traitement  $T$ , alors que l'effet de résultat estime l'impact de  $U$  sur la probabilité de l'issue  $Y$  lorsque  $T = 0$ .

On remarque que malgré des effets de sélection et de résultat non-négligeable, nos estimés demeurent cohérents avec ceux obtenus par nos régressions précédentes et n'infirmes pas nos résultats. Notons également que l'effet de sélection est toujours plus important que celui de résultat (1,92 à 2,67 pour l'effet de sélection contre 1,10 à 1,76 pour l'effet de résultat), ce qui explique probablement en grande partie pourquoi l'ATT demeure relativement inchangé.

**Tableau 4.5** Simulation de variable confondante : mathématiques

Statistique de sélection	Statistique de résultat	ATT	Écart-type	Effet de sélection	Effet de résultat
Médiane	Médiane	23,96***	2,27	2,07	1,76
Médiane	4e quartile	24,99***	2,21	1,94	1,48
4e quartile	Médiane	24,67***	2,30	2,61	1,29
4e quartile	4e quartile	25,56***	2,41	2,67	1,10

p-value : \* < 0,10; \*\* < 0,05; \*\*\* < 0,01

**Tableau 4.6** Simulation de variable confondante : lecture

Statistique de sélection	Statistique de résultat	ATT	Écart-type	Effet de sélection	Effet de résultat
Médiane	Médiane	24,76***	2,30	1,92	1,55
Médiane	4e quartile	24,57***	2,23	1,97	1,70
4e quartile	Médiane	24,98***	2,34	2,37	1,28
4e quartile	4e quartile	24,94***	2,40	2,55	1,29

p-value : \* < 0,10; \*\* < 0,05; \*\*\* < 0,01

**Tableau 4.7** Simulation de variable confondante : science

Statistique de sélection	Statistique de résultat	ATT	Écart-type	Effet de sélection	Effet de résultat
Médiane	Médiane	23,27***	2,36	2,16	1,46
Médiane	4e quartile	22,75***	2,33	1,96	1,70
4e quartile	Médiane	23,14***	2,45	2,33	1,36
4e quartile	4e quartile	22,10***	2,40	2,15	1,71

p-value : \* < 0,10 ; \*\* < 0,05 ; \*\*\* < 0,01

## CHAPITRE V

### DISCUSSION DES RÉSULTATS

Initialement, lorsque les écarts des résultats PISA des élèves du privé et ceux du public étaient observés de façon simpliste, on remarquait une différence allant de 54,21 % à 62,49 % d'un écart-type en faveur des élèves des établissements privés. Une fois l'effet de traitement estimé de façon rigoureuse, la différence passe à 29,02 % d'un écart-type dans le cas des scores en science, ou 34,1% dans le cas des mathématiques. Le tableau 5.1 illustre ceci en détails. Il est pertinent de noter que l'ATT demeure responsable de plus de la moitié de l'écart initialement observé, peu importe la matière concernée.

**Tableau 5.1** Scores PISA : écarts entre public et privé (pourcentages d'écart-type)

	Écart initial	ATT estimé	Pourcentage de l'écart expliqué
Mathématiques	62,49 %	34,1 %	54,57 %
Lecture	57,87 %	31,6 %	54,61 %
Science	54,21 %	29,02 %	53,53 %

Ces résultats concordent avec ceux obtenus par Lefebvre *et al.*, qui estimaient cet effet à 12 % à 36 % d'un écart-type (Lefebvre *et al.*, 2011). Il est également pertinent de comparer ces estimations avec celles effectuées par Angrist *et al* pour

les élèves des *charter schools* du Massachussets, qui obtenaient un effet de 36,7 % d'un écart-type en mathématiques et de 25,9 % à 33,4 % d'un écart-type pour la maîtrise de la langue anglaise (variant selon le type de compétence mesuré) pour les élèves d'âge comparable (Angrist *et al.*, 2011). L'assignation quasi-aléatoire à ces écoles en fait un élément de comparaison intéressant étant donné que le revenu des parents ou leur statut socio-économique n'a aucun impact sur l'admissibilité de l'élève à ces écoles. Or, une fois ces facteurs contrôlés dans le cas du Québec, nous obtenons une estimation d'ATT similaire. Il est donc raisonnable d'affirmer que des facteurs communs aux deux types d'écoles influencent les résultats des élèves du Québec et du Massachussets.

De plus, lorsque les résultats obtenus sont validés par la simulation de variable confondante sur la base des données du revenu familial disponibles pour la première enquête, l'ATT qui est ainsi ré-estimé ne diverge essentiellement pas des mesures initiales. Il apparaît difficile de concevoir une autre variable omise dont les effets de sélection et de résultat seraient plus importants que ceux du revenu familial, de telle façon que l'ATT ré-estimé serait nul ou significativement réduit. Les résultats apparaissent donc comme étant robustes aux effets de variables confondantes qui seraient potentiellement omises dans le modèle.

## CONCLUSION

Il apparaît clair que la fréquentation d'un établissement d'enseignement secondaire privé au Québec a un impact positif sur les traités, d'au moins 29,02 % d'un écart-type en moyenne mais pouvant aller jusqu'à 37,48 %, et qui est statistiquement significatif à 99 % selon toutes les estimations effectuées. Ces estimations de l'effet du traitement prennent en compte les biais de sélection et de causalité liées aux facteurs familiaux, socio-économiques et autres caractéristiques de l'élève et de son environnement qui pourraient être reliés au choix de la fréquentation de ce type d'établissement ainsi qu'à une performance académique supérieure. Elles semblent également robustes à l'omission de variables liées à la sélection du traitement et au résultat positif tels que mesurés par les scores aux tests PISA.

Toutefois, il est pertinent de s'interroger sur la façon dont se traduit cette différence dans les scores sur d'autres variables d'intérêt reliées à la performance scolaire, notamment l'obtention d'un diplôme d'éducation secondaire (DES) dans les délais normaux, c'est-à-dire au terme de la cinquième année du secondaire, sans redoublement. Les cycles ultérieurs de la cohorte A de l'EJET, suivant les individus du premier cycle (ayant fait le test PISA en 2000) jusqu'à l'âge de 25 ans, offrent plusieurs possibilités d'exploitation pour des études futures à ce niveau. Le modèle employé ici pourrait être reproduit, avec l'obtention du DES à 17 ans comme variable dépendante.

Une autre piste pour des recherches futures est l'identification des facteurs qui se trouvent à la source de l'effet de traitement. Plusieurs hypothèses ont été émises à cet effet et ont été discutées au premier chapitre, notamment les effets de pairs et

la qualité des enseignants. Étudier ce qui différencie à la fois les écoles à charte du Massachusetts et les écoles privées du Québec des systèmes publics avec lesquels ils compétitionnent pourrait donner d'autres pistes.

À cet effet, Fryer isole cinq facteurs pouvant expliquer le succès des écoles à charte américaines qui performant bien : elles mettent l'accent sur le capital humain de l'école (enseignants et directeurs), elles assurent un meilleur suivi de la performance des élèves et exploitent ces données afin de personnaliser l'enseignement qui leur est imparti, elles augmentent les heures de tutorat, donnent davantage d'heures d'enseignement, et finalement elles cherchent à développer une culture de hautes attentes (Fryer, 2012). La recherche future en éducation a de nombreux candidats comme facteurs explicatifs de l'effet de traitement des écoles privées du Québec.

## APPENDICE A

### TABLEAUX SUPPLÉMENTAIRES

Tableau A.1 Statistiques générales, 1/3

	Public %	Public n	Privé %	Privé n	Tous %	Tous n
Type d'école	79.82	13614	20.18	3442		17056
Année						
	2000	2982	18.42	673	21.43	3655
	2003	2405	17.03	494	17.00	2899
	2006	2779	19.12	657	20.15	3436
	2009	2593	21.90	727	19.47	3320
	2012	2849	23.94	897	21.96	3746
Ville	70.81	4071	29.19	1678	33.71	5750
Campagne	84.40	9542	15.60	1764	66.29	11306
Parents immigrants						
	0	10924	19.05	2572	79.13	13496
	1	966	25.65	333	7.62	1300
	2	1723	23.78	537	13.25	2260
Quintile SSE						
	1	3174	7.01	239	20.01	3413
	2	2961	13.17	449	19.99	3410
	3	2852	16.42	560	20.01	3412
	4	2508	26.44	902	19.99	3410
	5	2119	37.89	1292	20.00	3411

Source : Calcul de l'auteur à partir des données pondérées de PISA

Tableau A.2 Statistiques générales, 2/3

		Public %	Public n	Privé %	Privé n	Tous %	Tous n
Livres à la maison	0 / Données manquantes	87.57	255	12.43	36	1.71	291
	1-10	90.55	1604	9.45	167	10.39	1772
	11-100	83.80	6932	16.20	1340	48.50	8272
	101-500	72.65	4065	27.35	1531	32.81	5596
	500+	67.27	757	32.73	368	6.60	1126
Éducation de la mère (2000)	Données manquantes	83.69	591	16.31	115	19.32	706
	Sec. non complété	93.19	416	6.81	30	12.20	446
	Sec. complété	86.50	897	13.50	140	28.38	1037
	Tech. / prof. / collégial	80.35	679	19.65	166	23.13	845
	Uni. 1er cycle	67.33	331	32.67	161	13.47	492
	Uni. cycles sups.	52.53	67	47.47	61	3.50	128

Source : Calcul de l'auteur à partir des données pondérées de PISA

Tableau A.3 Statistiques générales, 3/3

	Public %	Public n	Privé %	Privé n	Tous %	Tous n
Éducation de la mère	91.24	202	8.76	19	1.66	222
(Post-2000)						
Données manquantes						
Sec. non complété	93.85	1007	6.15	66	8.00	1072
Sec. complété	88.90	1918	11.10	239	16.10	2157
Dip. professionnel	87.95	2103	12.05	288	17.84	2391
Dip. collégial	79.62	2897	20.38	741	27.15	3638
Dip. universitaire	64.16	2515	35.84	1405	29.25	3920
Langue parlée à la maison						
Français	79.51	10967	20.49	2826	80.87	13793
Anglais	79.39	1567	20.61	407	11.57	1974
Autre	83.73	1079	16.27	210	7.56	1289
Âge						
15	79.34	7345	20.66	1912	54.27	9257
16	80.38	6269	19.62	1530	45.73	7799
Niveau scolaire						
3	85.45	5172	14.55	881	35.49	6053
4	76.72	8442	23.28	2561	64.51	11003

Source : Calcul de l'auteur à partir des données pondérées de PISA

**Tableau A.4** Répartition des sexes selon le type d'établissement

	n	% de la sous-catégorie	% total
<b>Filles</b>			
Public	7009	79,60	41,09
Privé	1796	20,40	10,53
<b>Total</b>	<b>8805</b>	<b>100</b>	<b>51,62</b>
<b>Garçons</b>			
Public	6605	80,05	38,73
Privé	1646	19,95	9,65
<b>Total</b>	<b>8251</b>	<b>100</b>	<b>48,38</b>
<b>Observations</b>	<b>17056</b>		

Source : Calcul de l'auteur à partir des données pondérées de PISA

**Tableau A.5** Types d'établissement selon la taille des villes

Taille de la ville	n	%
<i>Moins de 100 000 habitants</i>	<i>11306</i>	<i>66,29</i>
Public	9542	84,4
Privé	1764	15,6
<i>100 000 habitants ou plus</i>	<i>5750</i>	<i>33,71</i>
Public	4072	70,81
Privé	1678	29,19
<b>Observations</b>	<b>17056</b>	

Source : Calcul de l'auteur à partir des données pondérées de PISA

**Tableau A.6** Nombre de parents immigrants selon le type d'établissement

	Aucun		Un seul parent		Deux parents	
	n	%	n	%	n	%
Public	10925	80,95	966	74,35	1723	76,22
Privé	2572	19,05	333	25,65	537	23,78
Total	13497	79,13	1299	7,62	2260	13,25
Observations	17056					

Source : Calcul de l'auteur à partir des données pondérées de PISA

**Tableau A.7** Langue parlée à la maison selon le type d'établissement

	n	%
<i>Français</i>	<i>13793</i>	<i>80,87</i>
Public	10967	79,51
Privé	2826	20,49
<i>Anglais</i>	<i>1974</i>	<i>11,57</i>
Public	1567	79,39
Privé	407	20,61
<i>Autre</i>	<i>1289</i>	<i>7,56</i>
Public	1079	83,73
Privé	210	16,27
Observations	17056	

Source : Calcul de l'auteur à partir des données pondérées de PISA

Tableau A.8 Score en mathématiques

Variable	Coefficient	Écart-type	Statistique t
Privé	26,70***	3,10	8,61
Garçon	18,30***	1,64	11,14
Un parent immigrant	-13,17***	2,99	-4,40
Deux parents immigrants	-23,44***	2,87	-8,17
2003	-17,73*	10,19	-1,74
2006	-23,68**	10,09	-2,35
2009	-17,98*	10,05	-1,79
2012	-27,98***	10,09	-2,77
2e quintile SSE	5,25	3,26	1,61
3e quintile SSE	11,13***	3,18	3,50
4e quintile SSE	13,45***	3,20	4,21
5e quintile SSE	25,31***	3,16	8,00
1-10 livres à la maison	-19,27**	8,59	-2,24
11-100 livres à la maison	3,92	8,04	0,49
101-500 livres à la maison	27,19***	8,01	3,39
>500 livres à la maison	37,94***	8,40	4,52
2000, secondaire non complété	-5,53	8,69	-0,64
2000, secondaire complété	4,63	5,66	0,82
2000, diplôme technique/professionnel/collégial	1,63	5,82	0,28
2000, universitaire 1er cycle	4,42	6,27	0,71
2000, universitaire cycles supérieurs	18,61*	10,70	1,74
Post-2000, secondaire non complété	16,30	10,34	1,58
Post-2000, secondaire complété	17,50*	9,22	1,90
Post-2000, diplôme professionnel	14,74	9,26	1,59
Post-2000, diplôme collégial	19,64**	9,06	2,17
Post-2000, diplôme universitaire	27,59***	9,07	3,04
Anglais parlé à la maison	-19,44***	2,10	-9,25
Autre langue parlée à la maison	-7,26*	4,19	-1,73
Élève d'au moins 16 ans	-7,99***	1,77	-4,51
Élève de secondaire 4	50,55***	2,04	24,73
Constante	497,59***	9,46	52,62

n=15412, R<sup>2</sup>=0,24, p-value : \*<0,10; \*\*<0,05; \*\*\*<0,01

Tableau A.9 Score en lecture

Variable	Coefficient	Écart-type	Statistique t
Privé	25,61***	3,87	6,63
Garçon	-24,83***	1,62	-15,30
Un parent immigrant	-8,49***	3,05	-2,78
Deux parents immigrants	-20,16***	2,83	-7,13
2003	-39,31***	10,91	-3,60
2006	-41,14***	10,92	-3,77
2009	-50,06***	10,82	-4,63
2012	-49,30***	10,84	-4,55
2e quintile SSE	4,41	3,22	1,37
3e quintile SSE	10,71***	3,01	3,55
4e quintile SSE	11,53***	3,11	3,71
5e quintile SSE	22,86***	3,04	7,52
1-10 livres à la maison	-15,85*	8,50	-1,86
11-100 livres à la maison	12,53	7,97	1,57
101-500 livres à la maison	35,38***	7,96	4,45
>500 livres à la maison	47,35***	8,35	5,67
2000, secondaire non complété	-3,60	7,96	-0,45
2000, secondaire complété	-1,90	5,05	-0,38
2000, diplôme technique/professionnel/collégial	-0,78	4,95	-0,16
2000, universitaire 1er cycle	16,38***	5,23	3,13
2000, universitaire cycles supérieurs	16,43*	8,75	1,88
Post-2000, secondaire non complété	38,51***	10,98	3,51
Post-2000, secondaire complété	35,65***	10,33	3,45
Post-2000, diplôme professionnel	32,71***	10,29	3,18
Post-2000, diplôme collégial	38,05***	10,15	3,75
Post-2000, diplôme universitaire	41,36***	10,14	4,08
Anglais parlé à la maison	-9,89***	2,11	-4,69
Autre langue parlée à la maison	-9,51**	4,25	-2,24
Élève d'au moins 16 ans	-10,67***	1,76	-6,05
Élève de secondaire 4	43,08***	2,03	21,25
Constante	506,34***	9,19	55,12

n=17056, R<sup>2</sup>=0,22, p-value : \*<0,10; \*\*<0,05; \*\*\*<0,01

Tableau A.10 Score en science

Variable	Coefficient	Écart-type	Statistique t
Privé	23,71***	3,49	6,80
Garçon	11,54***	1,77	6,54
Un parent immigrant	-11,44***	3,08	-3,72
Deux parents immigrants	-30,55***	3,44	-8,89
2003	-34,25***	11,57	-2,96
2006	-36,40***	11,50	-3,16
2009	-47,55***	11,44	-4,16
2012	-59,98***	11,46	-5,24
2e quintile SSE	5,29	3,80	1,39
3e quintile SSE	14,03***	3,61	3,89
4e quintile SSE	14,08***	3,59	3,92
5e quintile SSE	25,36***	3,56	7,11
1-10 livres à la maison	-21,11**	9,53	-2,21
11-100 livres à la maison	7,02	8,85	0,79
101-500 livres à la maison	29,53***	8,84	3,34
>500 livres à la maison	40,87***	9,30	4,39
2000, secondaire non complété	-16,71	13,59	-1,23
2000, secondaire complété	-5,17	7,84	-0,66
2000, diplôme technique/professionnel/collégial	0,05	7,29	0,01
2000, universitaire 1er cycle	13,07*	7,87	1,66
2000, universitaire cycles supérieurs	2,74	14,85	0,18
Post-2000, secondaire non complété	29,42***	11,07	2,66
Post-2000, secondaire complété	26,37***	10,00	2,64
Post-2000, diplôme professionnel	26,59***	10,04	2,65
Post-2000, diplôme collégial	29,55***	9,83	3,01
Post-2000, diplôme universitaire	37,48***	9,84	3,81
Anglais parlé à la maison	-15,98***	2,29	-6,99
Autre langue parlée à la maison	-10,04**	4,49	-2,24
Élève d'au moins 16 ans	-10,95***	1,87	-5,84
Élève de secondaire 4	44,73***	2,27	19,75
Constante	500,07***	11,02	45,38

n=15418, R<sup>2</sup>=0,22, p-value : \*<0,10; \*\*<0,05; \*\*\*<0,01

Tableau A.11 Score des filles en mathématiques

Variable	Coefficient	Écart-type	Statistique t
Privé	28,48***	3,75	7,60
Un parent immigrant	-4,30	4,02	-1,07
Deux parents immigrants	-18,45***	3,78	-4,88
2003	-13,07	12,07	-1,08
2006	-21,30*	11,96	-1,78
2009	-16,69	11,87	-1,41
2012	-26,77**	11,94	-2,24
2e quintile SSE	0,51	4,28	0,12
3e quintile SSE	8,24*	4,25	1,94
4e quintile SSE	4,58	4,22	1,08
5e quintile SSE	22,91***	4,22	5,43
1-10 livres à la maison	-27,82**	11,81	-2,35
11-100 livres à la maison	-7,12	10,91	-0,65
101-500 livres à la maison	18,31*	10,86	1,69
>500 livres à la maison	28,73**	11,36	2,53
2000, secondaire non complété	-23,68***	9,15	-2,59
2000, secondaire complété	4,05	6,87	0,59
2000, diplôme technique/professionnel/collégial	0,49	7,66	0,06
2000, universitaire 1er cycle	6,36	7,82	0,81
2000, universitaire cycles supérieurs	10,22	15,57	0,66
Post-2000, secondaire non complété	3,84	12,31	0,31
Post-2000, secondaire complété	11,34	11,13	1,02
Post-2000, diplôme professionnel	5,39	11,15	0,48
Post-2000, diplôme collégial	13,10	10,90	1,20
Post-2000, diplôme universitaire	23,49**	10,90	2,15
Anglais parlé à la maison	-21,26***	2,84	-7,47
Autre langue parlée à la maison	-10,32*	5,58	-1,85
Élève d'au moins 16 ans	-5,53**	2,41	-2,30
Élève de secondaire 4	46,95***	2,87	16,33
Constante	513,93***	12,07	42,59

n=8130, R<sup>2</sup>=0,24, p-value : \* < 0,10; \*\* < 0,05; \*\*\* < 0,01

Tableau A.12 Score des filles en lecture

Variable	Coefficient	Écart-type	Statistique t
Privé	26,54***	4,19	6,34
Un parent immigrant	-2,04	3,79	-0,54
Deux parents immigrants	-18,25***	3,72	-4,91
2003	-32,77**	13,14	-2,49
2006	-38,24***	13,23	-2,89
2009	-40,28***	12,99	-3,10
2012	-42,17***	13,03	-3,24
2e quintile SSE	2,72	4,15	0,66
3e quintile SSE	9,07**	3,95	2,30
4e quintile SSE	4,51	4,05	1,11
5e quintile SSE	21,39***	4,01	5,34
1-10 livres à la maison	-18,74*	9,60	-1,95
11-100 livres à la maison	4,68	8,80	0,53
101-500 livres à la maison	30,82***	8,78	3,51
>500 livres à la maison	43,52***	9,36	4,65
2000, secondaire non complété	0,27	11,33	0,02
2000, secondaire complété	11,34*	6,75	1,68
2000, diplôme technique/professionnel/collégial	9,44	6,67	1,41
2000, universitaire 1er cycle	25,97***	7,02	3,70
2000, universitaire cycles supérieurs	23,81**	12,07	1,97
Post-2000, secondaire non complété	33,73**	13,19	2,56
Post-2000, secondaire complété	35,62***	12,23	2,91
Post-2000, diplôme professionnel	29,88**	12,20	2,45
Post-2000, diplôme collégial	36,78***	12,02	3,06
Post-2000, diplôme universitaire	42,85***	11,96	3,58
Anglais parlé à la maison	-12,18***	2,77	-4,39
Autre langue parlée à la maison	-7,23	5,45	-1,33
Élève d'au moins 16 ans	-8,19***	2,35	-3,49
Élève de secondaire 4	40,16***	2,80	14,35
Constante	507,45***	10,46	48,50

n=8971, R<sup>2</sup>=0,21, p-value : \*<0,10; \*\*<0,05; \*\*\*<0,01

Tableau A.13 Score des filles en science

Variable	Coefficient	Écart-type	Statistique t
Privé	25,05***	3,60	6,95
Un parent immigrant	-5,18	3,92	-1,32
Deux parents immigrants	-26,63***	4,29	-6,20
2003	-24,42	15,53	-1,57
2006	-29,07*	15,43	-1,88
2009	-38,34**	15,29	-2,51
2012	-49,49***	15,36	-3,22
2e quintile SSE	5,80	4,78	1,21
3e quintile SSE	10,96**	4,78	2,29
4e quintile SSE	6,92	4,73	1,46
5e quintile SSE	24,36***	4,74	5,14
1-10 livres à la maison	-17,10	12,91	-1,33
11-100 livres à la maison	1,42	12,07	0,12
101-500 livres à la maison	27,55**	12,05	2,29
>500 livres à la maison	39,67***	12,63	3,14
2000, secondaire non complété	-0,38	19,66	-0,02
2000, secondaire complété	17,12	10,55	1,62
2000, diplôme technique/professionnel/collégial	9,68	10,81	0,89
2000, universitaire 1er cycle	30,96***	11,21	2,76
2000, universitaire cycles supérieurs	32,77**	14,42	2,27
Post-2000, secondaire non complété	22,95*	13,89	1,65
Post-2000, secondaire complété	25,26**	12,67	1,99
Post-2000, diplôme professionnel	23,22*	12,75	1,82
Post-2000, diplôme collégial	27,12**	12,47	2,17
Post-2000, diplôme universitaire	36,50***	12,47	2,93
Anglais parlé à la maison	-17,50***	3,00	-5,84
Autre langue parlée à la maison	-12,27**	5,84	-2,10
Élève d'au moins 16 ans	-9,44***	2,49	-3,79
Élève de secondaire 4	41,31***	3,02	13,67
Constante	497,41***	15,06	33,03

n=8130, R<sup>2</sup>=0,23, p-value : \*<0,10; \*\*<0,05; \*\*\*<0,01

Tableau A.14 Score des garçons en mathématiques

Variable	Coefficient	Écart-type	Statistique t
Privé	24,78***	3,94	6,28
Un parent immigrant	-23,60***	4,47	-5,28
Deux parents immigrants	-29,79***	4,37	-6,82
2003	-29,04*	16,09	-1,80
2006	-30,88*	15,90	-1,94
2009	-24,54	15,92	-1,54
2012	-34,78**	15,95	-2,18
2e quintile SSE	10,72**	4,88	2,20
3e quintile SSE	14,45***	4,69	3,08
4e quintile SSE	22,81***	4,74	4,81
5e quintile SSE	27,88***	4,65	6,00
1-10 livres à la maison	-9,21	12,26	-0,75
11-100 livres à la maison	16,16	11,56	1,40
101-500 livres à la maison	37,01***	11,53	3,21
>500 livres à la maison	47,64***	12,15	3,92
2000, secondaire non complété	10,68	13,20	0,81
2000, secondaire complété	9,39	9,49	0,99
2000, diplôme technique/professionnel/collégial	7,28	9,31	0,78
2000, universitaire 1er cycle	6,50	10,09	0,64
2000, universitaire cycles supérieurs	30,21**	14,16	2,13
Post-2000, secondaire non complété	39,48**	16,05	2,46
Post-2000, secondaire complété	31,99**	14,02	2,28
Post-2000, diplôme professionnel	33,59**	14,08	2,39
Post-2000, diplôme collégial	34,34**	13,79	2,49
Post-2000, diplôme universitaire	40,05***	13,76	2,91
Anglais parlé à la maison	-16,93***	3,10	-5,46
Autre langue parlée à la maison	-3,95	6,33	-0,62
Élève d'au moins 16 ans	-10,46***	2,61	-4,01
Élève de secondaire 4	54,29***	2,90	18,71
Constante	495,43***	14,67	33,76

n=7282, R<sup>2</sup>=0,23, p-value : \*<0,10; \*\*<0,05; \*\*\*<0,01

Tableau A.15 Score des garçons en lecture

Variable	Coefficient	Écart-type	Statistique t
Privé	24,52***	5,35	4,59
Un parent immigrant	-15,49***	4,95	-3,13
Deux parents immigrants	-22,56***	4,33	-5,21
2003	-51,92***	16,30	-3,19
2006	-49,41***	16,20	-3,05
2009	-65,41***	16,18	-4,04
2012	-62,35***	16,22	-3,84
2e quintile SSE	7,12	4,90	1,45
3e quintile SSE	12,90***	4,53	2,85
4e quintile SSE	19,34***	4,67	4,14
5e quintile SSE	25,06***	4,54	5,52
1-10 livres à la maison	-11,68	14,39	-0,81
11-100 livres à la maison	21,33	13,75	1,55
101-500 livres à la maison	40,79***	13,74	2,97
>500 livres à la maison	51,49***	14,32	3,60
2000, secondaire non complété	-6,39	11,00	-0,58
2000, secondaire complété	-12,33*	7,22	-1,71
2000, diplôme technique/professionnel/collégial	-9,86	7,10	-1,39
2000, universitaire 1er cycle	7,38	7,67	0,96
2000, universitaire cycles supérieurs	11,92	11,93	1,00
Post-2000, secondaire non complété	49,94***	16,57	3,01
Post-2000, secondaire complété	40,85***	15,57	2,62
Post-2000, diplôme professionnel	42,52***	15,45	2,75
Post-2000, diplôme collégial	45,01***	15,22	2,96
Post-2000, diplôme universitaire	45,67***	15,22	3,00
Anglais parlé à la maison	-7,17**	3,26	-2,20
Autre langue parlée à la maison	-11,75*	6,66	-1,76
Élève d'au moins 16 ans	-12,78***	2,63	-4,85
Élève de secondaire 4	45,68***	2,90	15,75
Constante	478,72***	15,41	31,07

n=8085, R<sup>2</sup>=0,19, p-value : \*<0,10; \*\*<0,05; \*\*\*<0,01

Tableau A.16 Score des garçons en science

Variable	Coefficient	Écart-type	Statistique t
Privé	21,93***	4,83	4,54
Un parent immigrant	-19,24***	4,85	-3,97
Deux parents immigrants	-34,31***	4,90	-7,01
2003	-49,16***	15,72	-3,13
2006	-47,99***	15,66	-3,06
2009	-60,77***	15,62	-3,89
2012	-75,22***	15,63	-4,81
2e quintile SSE	6,07	5,72	1,06
3e quintile SSE	18,13***	5,26	3,44
4e quintile SSE	22,48***	5,26	4,27
5e quintile SSE	28,07***	5,18	5,41
1-10 livres à la maison	-23,63*	13,91	-1,70
11-100 livres à la maison	12,98	13,02	1,00
101-500 livres à la maison	31,99**	13,02	2,46
>500 livres à la maison	42,46***	13,74	3,09
2000, secondaire non complété	-23,33	17,84	-1,31
2000, secondaire complété	-19,50*	10,79	-1,81
2000, diplôme technique/professionnel/collégial	-6,08	9,83	-0,62
2000, universitaire 1er cycle	-3,03	11,17	-0,27
2000, universitaire cycles supérieurs	-16,28	21,52	-0,76
Post-2000, secondaire non complété	43,89***	15,77	2,78
Post-2000, secondaire complété	34,02**	13,85	2,46
Post-2000, diplôme professionnel	38,59***	13,83	2,79
Post-2000, diplôme collégial	39,18***	13,50	2,90
Post-2000, diplôme universitaire	45,40***	13,54	3,35
Anglais parlé à la maison	-13,78***	3,46	-3,98
Autre langue parlée à la maison	-8,28	6,70	-1,24
Élève d'au moins 16 ans	-11,97***	2,80	-4,28
Élève de secondaire 4	46,84***	3,18	14,75
Constante	510,74***	16,05	31,83

n=7288, R<sup>2</sup>=0,21, p-value : \*<0,10; \*\*<0,05; \*\*\*<0,01

## BIBLIOGRAPHIE

- Altonji, J. G., Huang, C.-I. et Taber, C. R. (2010). Estimating the Cream Skimming Effect of School Choice. *National Bureau of Economic Research Working Paper Series, No. 16579(2)*, 266–324.
- Angrist, J. D., Cohodes, S. R., Dynarski, S. M., Fullerton, J. B., Kane, T. J., Pathak, P. A. et Walters, C. R. (2011). Student Achievement in Massachusetts' Charter Schools. *Center for Education Policy Research, Harvard University*.
- Booker, K., Gilpatric, S. M., Gronberg, T. et Jansen, D. (2008). The effect of charter schools on traditional public school students in Texas : Are children who stay behind left behind? *Journal of Urban Economics*, 64(1), 123–145.
- Boucher, V., Bramoullé, Y., Djebbari, H. et Fortin, B. (2014). Do peers affect student achievement? Evidence from Canada using group size variation. *Journal of Applied Econometrics*, 29(1), 91–109.
- Carnroy, M. et Rothstein, R. (2013). What Do International Tests Really Show about U.S. Student Performance. *Economic Policy Institute*.
- Frenette, M., Ching, P. et Chan, W. (2015). Analytical Studies Branch Research Paper Series Academic Outcomes of Public and Private High School Students : What Lies Behind the Differences? *Analytic Studies Branch Research Paper Series*, (367).
- Fryer, R. G. J. (2012). Learning from the successes and failures of charter schools. *The Hamilton Project, Brookings*, (2012-06).
- Ganzeboom, H. et Treiman, D. (1996). Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Social science research*, 239(25), 201–239.
- Hainmueller, J. (2012). Entropy balancing for causal effects : A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1), 25–46.
- Hanushek, E. A. et Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*, 46(3), 607–668.

- Ichino, A., Mealli, F. et Nannicini, T. (2008). From temporary help jobs to permanent employment : What can we learn from matching estimators and their sensitivity? *Journal of Applied Econometrics*, 23(3), 305–327.
- Imberman, S. a. (2011). The effect of charter schools on achievement and behavior of public school students. *Journal of Public Economics*, 95(7-8), 850–863.
- Lalonde, R. J. (1986). Evaluating the Econometric Evaluations of Training Programs with Experimental Data. *The American Economic Review*, 76(4), 604–620.
- Lefebvre, P., Merrigan, P. et Verstraete, M. (2011). Public subsidies to private schools do make a difference for achievement in mathematics : Longitudinal evidence from Canada. *Economics of Education Review*, 30(1), 79–98.
- McMullen, K. et Gilmore, J. (2010). A Note on High School Graduation and School Attendance, by Age and Province, 2009/2010. *Statistics Canada. Education Matters : Insights on Education, Learning and Training in Canada*, 7(4).
- Murnane, R., Willett, J., Duhaldeborde, Y. et Tyler, J. (2000). How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management*, 19(4), 547–568.
- Murnane, R. J., Willett, J. B. et Levy, F. (1995). The Growing Importance of Cognitive Skills in Wage Determination. *Review of Economics and Statistics*, 77(2), 251–266.
- Ni, Y. (2009). The impact of charter schools on the efficiency of traditional public schools : Evidence from Michigan. *Economics of Education Review*, 28(5), 571–584.
- OCDE. (2009). *PISA Data Analysis : SAS Second Edition*. OECD Publishing.
- OCDE. (2012). *Public and private schools : How management and funding relate to their socio-economics profile*. OECD Publishing.
- Vandenberghe, V. et Robin, S. (2004). Evaluating the effectiveness of private education across countries : A comparison of methods. *Labour Economics*, 11(4), 487–506.