

Dominique Dionne-Simard (6613890)

Analyse de l'écart entre les élèves d'écoles de langue minoritaire et d'écoles de langue majoritaire
sur les résultats aux tests d'aptitudes PISA de compréhension en lecture et de raisonnement
mathématique : Une application de la décomposition Blinder-Oaxaca

Mémoire supervisé par Mme Fernanda Estevan, Ph.D. et présenté au
département de science économique de l'Université d'Ottawa

Département de science économique

Université d'Ottawa

Le 14 janvier 2013

Table des matières

1. Introduction	3
2. Revue de littérature	5
3. L'étude PISA	13
4. Données canadiennes et variables d'intérêts	16
5. Méthode empirique	19
6. Résultats	22
<i>6.1 Les fonctions de production d'éducation</i>	22
6.1.1 Compréhension en lecture	22
6.1.2 Raisonnement mathématique	24
<i>6.2 Les décompositions Blinder-Oaxaca</i>	26
6.2.1 Compréhension en lecture	26
6.2.2 Raisonnement mathématique	28
<i>6.3 Les décompositions Blinder-Oaxaca pour l'Ontario, l'Alberta et la Colombie-Britannique</i>	29
7. Discussion	29
8. Conclusion	34
Références	37
Annexe 1: Description des variables d'intérêts	41
Tableaux	45

1. Introduction

En 1982, le Canada adoptait la *Charte canadienne des droits et libertés* qui constitue la partie 1 de la *Loi constitutionnelle de 1982*. Compris dans la Charte, l'article 23 garantit aux minorités linguistiques francophones et anglophones du Canada, là où le nombre le justifie, le droit à une éducation publique dans la langue de la minorité (*Charte canadienne des droits et libertés*, 1982). Dans l'arrêt *Arseneault-Cameron c Île-du-Prince-Édouard*, la Cour Suprême du Canada a déclaré que : « L'article 23 repose sur la prémisse que l'égalité réelle exige que les minorités de langue officielle soient traitées différemment, si nécessaire, suivant leur situation et leurs besoins particuliers, afin de leur assurer un niveau d'éducation équivalent à celui de la majorité de langue officielle » (*Arseneault-Cameron c Île-du-Prince-Édouard*, 2000 : para 31).

À la lumière de l'obligation constitutionnelle que représente l'article 23 de la Charte, il convient d'analyser dans quelle mesure cette obligation est respectée à l'intérieur des provinces canadiennes. Le présent mémoire est motivé par les questions suivantes : Les élèves qui fréquentent des écoles de langue minoritaire au Canada jouissent-ils d'une égalité réelle en matière de prestation de services dans la langue de la minorité conformément à l'article 23 de la Charte? Dans la mesure où la réponse à la question précédente est négative, qu'est-ce qui conditionne l'écart entre les élèves d'écoles de langue minoritaire et les élèves d'écoles de langue majoritaire? Nous entendons par écoles de langue minoritaire les écoles de langues officielles minoritaires, c'est-à-dire le français dans le Canada hors Québec et l'anglais au Québec. De la même façon, les écoles de langue majoritaire du Canada sont celles de langues officielles, soit l'anglais dans le Canada hors Québec et le français au Québec. Les élèves ont été regroupés selon la langue d'enseignement de l'école qu'ils fréquentent.

Afin de répondre aux questions qui motivent ce mémoire, nous utilisons conjointement des fonctions de production d'éducation et la méthode Blinder-Oaxaca pour mesurer et décomposer l'écart des habiletés cognitives entre les élèves qui fréquentent une école de langue minoritaire et ceux qui fréquentent une école de langue majoritaire. Dans la mesure où le développement des habiletés cognitives est au centre du processus éducatif, la qualité de la prestation de services peut être mesurée en termes du niveau des habiletés cognitives. Nous adoptons à cet effet une analyse économique du processus éducatif (Hanushek, 1986). Nous partons de la prémisse que le niveau d'aptitudes mesuré en termes de résultats scolaires est fonction d'un ensemble d'intrants individuels, familiaux et institutionnels. Le processus éducatif peut ainsi être vu comme un investissement en capital humain. Nous utilisons les

résultats des tests d'aptitudes en littératie PISA (*Programme for International Student Assessment*) de l'OCDE (Organisation de coopération et de développement économiques) afin de mesurer l'écart d'habiletés cognitives entre les élèves d'écoles de langue minoritaire et les élèves d'écoles de langue majoritaire du Canada. Comme le souligne Hanushek (1986), le résultat de tests standardisés, bien que constituant une mesure imparfaite du niveau réel d'habiletés des élèves, est la mesure d'extrants la plus utilisée dans la littérature en économie de l'éducation. Les tests standardisés sont particulièrement pertinents lorsqu'il s'agit de mesurer des habiletés cognitives de base telles que la compréhension en lecture ou le raisonnement mathématique. Ainsi pour notre analyse, nous utilisons les résultats des tests d'aptitudes PISA de compréhension en lecture et de raisonnement mathématique. Nous comparons les deux groupes, les élèves d'écoles de langue minoritaire et les élèves d'écoles de langue majoritaire, à l'aide de la méthode Blinder-Oaxaca (Oaxaca, 1973 ; Blinder, 1973) qui nous permet de décomposer l'écart entre groupes en une partie expliquée par des différences observées au niveau des variables explicatives de la fonction de production éducative et une autre partie inexpliquée. Cette partie inexpliquée capture l'effet des retours sur les caractéristiques en termes de résultats ainsi que l'ensemble des variables qui ne sont pas incluses dans le modèle.

Nous posons l'hypothèse que les minorités linguistiques du Canada font face à des difficultés particulières en matière de construction des habiletés cognitives en raison de leur statut de minorité linguistique et que ces difficultés se retrouvent dans l'écart inexpliqué. En effet, le manque d'accès à des ressources éducatives dans la langue de la minorité et l'exposition régulière à la langue de la majorité sont autant de facteurs qui peuvent influencer les résultats des élèves d'écoles de langue minoritaire sans pour autant affecter les élèves d'écoles de langue majoritaire. Breton (1999) parle de pressions assimilatrices et de dérive linguistique pour décrire la situation particulière des minorités linguistiques. Nous qualifions d'effet minoritaire l'ensemble des difficultés propres aux minorités linguistiques.

Les résultats montrent qu'il existe un écart statistiquement significatif en faveur des majorités linguistiques en ce qui a trait aux résultats sur les tests d'aptitudes (sauf sur le test de compréhension en lecture pour les élèves du Québec), qu'une partie de l'écart reste inexpliqué après l'avoir conditionné sur un ensemble de variables explicatives et que cet écart inexpliqué défavorise autant les élèves d'écoles de langue minoritaire francophone au Canada hors Québec qu'anglophone au Québec. Les résultats montrent aussi que l'écart inexpliqué sur le test de

compréhension en lecture représente une plus grande proportion de l'écart observé que l'écart inexpliqué sur le test de raisonnement mathématique au Canada hors Québec. Finalement, les résultats montrent que l'écart inexpliqué est plus important dans les provinces où les membres de la minorité linguistique sont moins nombreux et moins concentrés géographiquement (ex. Alberta et Colombie-Britannique) que dans les provinces où la minorité linguistique est plus nombreuse et davantage concentrée géographiquement (ex. Ontario).

Le présent mémoire de recherche est structuré comme suit. À la section 2, nous présentons un survol de la littérature en économie de l'éducation et plus particulièrement en ce qui a trait aux analyses d'écarts entre groupes. Aux sections 3 et 4 respectivement, nous présentons l'étude PISA et les données utilisées dans l'analyse et à la section 5, la stratégie empirique. Les résultats sont présentés à la section 6 et la section 7 propose une discussion de ces résultats. La section 8 conclut.

2. Revue de littérature

La présente revue de littérature débute par un survol de la littérature en économie de l'éducation, et plus particulièrement sur la littérature concernant les fonctions de production d'éducation. Nous poursuivons avec une revue de la littérature portant sur l'analyse d'écarts entre groupes en ce qui a trait aux fonctions de production d'éducation, et plus particulièrement sur l'utilisation de la méthode Blinder-Oaxaca en économie de l'éducation. Nous terminons avec un survol de la littérature dans le contexte canadien.

La genèse de l'analyse économique du processus éducationnel est souvent associée au rapport Coleman aux États-Unis: *Equality of Educational Opportunity* (Coleman *et al.*, 1966). Comme le mentionne Hanushek (1986), le mandat du rapport était de nature descriptive, visant essentiellement à produire un portrait des ressources éducationnelles dans les écoles du pays. Cependant, le rapport est allé beaucoup plus loin en produisant une base de données comprenant plus de 500 000 élèves répartis dans plus de 3000 écoles. Cette base de données comprenait, entre autres, des informations sur les résultats scolaires, sur le statut socio-économique et le statut ethnique des élèves. L'objectif du projet était de mesurer l'effet des ressources éducationnelles sur les résultats des élèves. Le rapport concluait notamment que les variables institutionnelles avaient peu d'effet sur les résultats. Toujours selon Hanushek (1986), les conclusions du rapport ont été amplement critiquées, autant dans la sphère politique que dans la sphère académique. Les

critiques du rapport Coleman ont cependant participé à l'essor d'un nouveau champ d'études en science économique, l'économie de l'éducation.

L'approche employée dans le rapport Coleman est encore aujourd'hui au centre de l'analyse économique de l'éducation. L'éducation y est perçue comme un processus de production intégrant un ensemble d'intrants afin de produire un extrant (Hanushek, 1986). Au niveau de l'élève, l'extrait est généralement mesuré en termes de résultats sur des tests standardisés afin de permettre la comparaison entre élèves. Sont compris dans les intrants, l'ensemble des caractéristiques individuelles (sexe, ethnicité, motivation, etc.), familiales (niveau d'éducation des parents, revenus des parents, matériel didactique disponible à la maison, etc.), et institutionnelles (qualité des enseignants, matériel didactique disponible à l'école, effet de pairs, etc.). Force est de constater que, bien qu'un ensemble d'intrants soient plus facilement identifiables et mesurables, un autre ensemble d'intrants, tels que la motivation des élèves et l'importance de l'éducation pour les parents, sont souvent absents des bases de données faute de mesures adéquates. De plus, la fonction de production d'éducation est une mesure instantanée d'un processus continu. En effet, les résultats mesurés lors de tests d'aptitudes standardisés sont fonction non seulement des intrants immédiats, mais aussi des intrants qui ont caractérisé le processus éducationnel de l'élève à travers le temps. L'interprétation des résultats découlant de la fonction de production d'éducation doit se faire à la lumière de ces particularités propres à la mesure du processus éducationnel.

Les résultats des études empiriques sur le processus éducationnel sont de deux ordres. Au niveau des caractéristiques individuelles et familiales, les résultats sont constants à travers la littérature (Hanushek, 1986).¹ La motivation de l'élève envers l'école, le revenu des parents et le niveau d'éducation des parents ont été associés positivement aux résultats des élèves. Quant aux intrants institutionnels, les associations restent ambiguës. Lee et Barro (2001) ont conclu que des classes plus petites et des salaires plus élevés pour les enseignants ont un effet positif sur les résultats des élèves. Woessmann et West (2006), en analysant l'effet de classe sur les résultats des élèves dans 11 pays européens, ont conclu que l'effet de classe n'est statistiquement significatif que pour deux des pays analysés et que cet effet n'est observé que dans les pays où le salaire des enseignants est relativement faible. La qualité des enseignants a été généralement

¹ Les résultats recensés par Hanushek (1986) sont cohérents avec les résultats des études empiriques sur la fonction de production d'éducation des années récentes. Voir notamment Lee et Barro (2001) pour des résultats plus récents.

associée de manière statistiquement significative aux résultats des élèves, mais ce qui constitue un enseignant efficace reste ambigu. Chingos et Peterson (2011) n'ont trouvé aucune corrélation entre le niveau d'éducation des enseignants et les résultats des élèves et Woessmann (2011) a conclu que les résultats des élèves sur des tests standardisés sont significativement plus élevés lorsque les enseignants sont rémunérés selon la performance de leurs élèves.

L'attrait important de la fonction de production d'éducation, au sens d'Hanushek (1986), réside notamment dans sa mesure de la qualité de l'éducation par opposition à la quantité d'éducation. La quantité réfère au nombre d'années d'éducation et la qualité réfère au processus de développement des habiletés cognitives. Plusieurs études ont d'ailleurs souligné le lien entre qualité d'éducation et revenus, et qualité d'éducation et croissance économique (Behrman et Birdsall, 1983 ; Hanushek et Kimko, 2000 ; Hanushek et Woessmann, 2008 ; Neal et Johnson, 1996). Behrman et Birdsall (1983) ont souligné l'importance des écarts dans la qualité d'éducation afin d'expliquer les écarts salariaux entre les régions du Brésil. Ils ont notamment trouvé que l'utilisation de la quantité d'éducation tend à surestimer l'écart au niveau des revenus entre les régions géographiques brésiliennes. Neal et Johnson (1996) ont conclu que l'écart salarial entre les noirs et les blancs aux États-Unis était en grande partie expliqué par un écart au niveau des habiletés découlant de la qualité de l'éducation reçue. Hanushek et Kimko (2000) ont par ailleurs conclu qu'il y a un lien statistiquement significatif entre les résultats obtenus sur des tests standardisés et la croissance économique d'un pays. L'extrait de la fonction de production d'éducation est donc aussi un intrant dans les fonctions de revenus et de croissance économique. Ainsi, un écart au niveau des habiletés cognitives découlant d'un écart au niveau de la qualité de l'éducation reçue peut avoir des conséquences défavorables à long terme pour un groupe ou un pays particulier.²

À la lumière des résultats des études empiriques sur la relation habiletés cognitives-revenus et habiletés cognitives-croissance, il n'est pas surprenant de constater que l'analyse de l'écart sur les tests standardisés entre groupes ou pays ait suscité un intérêt prononcé chez les chercheurs en économie de l'éducation. En effet, dans la mesure où les résultats sur les tests d'aptitudes standardisés ont été associés aux revenus et à la croissance, les écarts observés relativement aux résultats des fonctions de production éducative peuvent être étendus à la

² Voir aussi Hanushek et Luque (2003) pour un résumé exhaustif de la relation habiletés cognitives-revenus et habiletés cognitives-croissance retracée dans la littérature.

compréhension des inégalités en termes de revenus et des disparités internationales en termes de croissance économique.

L'analyse de l'écart entre groupes est aussi intéressante du point de vue méthodologique. Une des faiblesses de la fonction de production éducative, soulignée par Hanushek (1986), réside dans la difficulté à cerner l'effet causal de certains intrants sur les résultats.³ En effet, la nature des intrants dans le processus de production éducatif mène souvent à des problèmes d'endogénéité. Or, dans l'analyse de l'écart entre les groupes ou les pays, l'effet relatif des intrants est plus pertinent que leur effet absolu. Le problème d'endogénéité peut donc persister sans miner les conclusions relatives à l'analyse de l'écart dans la mesure où les biais sont constants à travers les groupes comparés (Hanushek et Woessmann, 2010).

L'analyse classique de l'écart entre pays repose sur une comparaison descriptive de l'effet des intrants dans la fonction de production d'éducation.⁴ Woessman (2003) a utilisé les données du *Trends in International Mathematics and Science Study* (TIMSS) afin de comparer les pays sur la base des effets relatifs des intrants dans la fonction de production agrégée au niveau international. L'échantillon de Woessmann (2003) comprenait plus de 260 000 élèves répartis dans 39 pays. Il a trouvé premièrement que les intrants individuels et familiaux avaient un effet statistiquement significatif sur les résultats des élèves. Notamment, les filles réussissaient moins bien que les garçons sur les tests d'aptitudes en mathématiques et en sciences. Le niveau de revenus des parents était corrélé positivement avec les résultats de l'élève. De plus, les élèves nés dans le pays où ils répondaient aux tests réussissaient mieux que ceux qui sont nés dans un autre pays. Au niveau des intrants institutionnels, Woessmann (2003) a trouvé que les résultats des élèves étaient positivement corrélés avec le niveau d'éducation des enseignants et avec le nombre d'ordinateurs dans l'école. Au-delà des ressources institutionnelles, les politiques institutionnelles avaient un effet encore plus important sur les résultats des élèves. Woessmann (2003) est arrivé à la conclusion que l'effet combiné des politiques institutionnelles telles que l'autonomie des enseignants, le statut privé ou public de l'école, et le partage des responsabilités à l'intérieur de l'école entre la direction et le corps enseignant, pouvait expliquer un écart de 2

³ Hanushek (1986) mentionne notamment que la difficulté d'interprétation de l'effet des intrants dans la forme simple d'une fonction de production éducative a grandement limité l'écoute des décideurs publics en matière d'économie de l'éducation.

⁴ Voir Hanushek et Woessmann (2010) pour un portrait exhaustif des analyses comparatives en économie de l'éducation.

écarts-types entre les élèves de différents pays. Il conclut que les différences observées entre les pays sont davantage une conséquence des politiques éducatives que des ressources disponibles.

Nonobstant la contribution de cette étude, la méthode employée par Woessmann (2003) demeure peu utile si l'objectif est de comparer directement les groupes entre eux. La méthode du modèle agrégé rend compte de l'effet de chaque intrant sur les résultats au niveau international sans pour autant permettre d'analyser l'écart observé entre les groupes et ce qui peut l'expliquer. À ce titre, l'utilisation de modèles agrégés avec variable discriminante pour le groupe est déjà plus appropriée. Schnepf (2007) a analysé les écarts sur les tests d'aptitudes PISA, TIMSS et PIRLS (*Program of International Reading Literacy Study*) entre les natifs et les immigrants de première et de deuxième génération dans 10 pays où la population immigrante représente plus de 10 % de la population totale. En conditionnant l'écart des résultats sur la langue parlée à la maison, Schnepf (2007) a trouvé que les résultats des étudiants sont en moyenne plus faibles lorsque la langue parlée à la maison n'est pas la langue du test, et ce dans tous les pays de l'étude. Dans 80 % des pays de langue anglaise, l'écart inexplicé (la valeur de l'estimateur de la variable discriminante) disparaissait suite au conditionnement sur la langue parlée à la maison, suggérant ainsi que l'écart sur les tests d'aptitudes entre immigrants et natifs était en fait dû à un écart au niveau des habiletés langagières. Lorsque l'écart était conditionné sur un ensemble de variables socio-économiques et sur la concentration d'immigrants dans l'école, l'écart entre immigrants et natifs était statistiquement significatif dans seulement 3 pays (France, Pays-Bas et Suisse). Meunier (2011), employant la méthode de la variable discriminante, est arrivé aux mêmes conclusions que Schnepf (2007) en ce qui a trait aux résultats des immigrants de première et de deuxième génération sur les tests d'aptitudes PISA en Suisse. Meunier (2011) a trouvé que l'écart était toujours significatif après conditionnement sur un ensemble d'intrants et que cet écart était plus important chez les immigrants de première génération.

Fryer et Levitt (2004) ont utilisé les données de l'*Early Childhood Longitudinal Study* (ECLS-K) au niveau de la maternelle afin de comparer les résultats entre les élèves noirs et les élèves blancs aux États-Unis. En utilisant la méthode de la variable discriminante et en contrôlant pour seulement six intrants, les auteurs ont trouvé que l'écart entre les élèves blancs et les élèves noirs n'était pas statistiquement significatif. Cependant, l'écart entre les élèves blancs et les élèves noirs prenait de l'ampleur au cours des deux premières années d'éducation, au point de devenir significatif. Les auteurs ont attribué cette amplification de l'écart à la qualité inférieure

des écoles fréquentées par les élèves noirs. Ainsi, une inégalité quant à la qualité des institutions d'enseignement serait à la base des écarts observés entre les élèves noirs et les élèves blancs au niveau des résultats scolaires à l'école secondaire.

La méthode de la variable discriminante a aussi été utilisée afin d'analyser l'écart sur des tests d'aptitudes standardisés entre autochtones et non autochtones en Australie (Leigh et Gong, 2009). Les auteurs ont utilisé les données du *Longitudinal Survey of Australian Children* afin d'analyser l'écart sur les tests d'aptitudes dans une cohorte d'enfants de 4 à 5 ans. Ils ont observé qu'après avoir contrôlé pour les intrants socio-économiques, l'écart inexplicé ne représentait plus qu'un à deux tiers de l'écart non conditionnel. Cependant, une partie statistiquement significative de l'écart restait inexplicée. À l'instar de Fryer et Levitt (2004), Leigh et Gong (2009) ont souligné que l'écart observé entre les autochtones et les non-autochtones tendait à prendre de l'ampleur avec le nombre d'années d'éducation, suggérant une asymétrie au niveau de la qualité de l'éducation entre les deux groupes.

L'utilisation de la méthode d'estimation de fonction de production d'éducation avec variable discriminante pour les groupes nous permet donc de mesurer, conditionnellement sur un ensemble d'intrants, l'effet d'appartenir à un groupe en termes de résultats. Il s'agit en fait de l'écart qui n'est pas expliqué par les intrants. Or, l'écart observé est en fait composé d'une part expliquée par les intrants et d'une part inexplicée par les intrants qui comprend les intrants non-mesurés et les retours des intrants en termes de résultats. La méthode Blinder-Oaxaca (Blinder, 1973 ; Oaxaca, 1973), développée dans le cadre de l'économie du travail afin de tester l'hypothèse de discrimination salariale entre les hommes et les femmes, est particulièrement pertinente dans le cadre d'analyses de l'économie de l'éducation. Intuitivement, il s'agit premièrement d'estimer le résultat du groupe A si celui-ci avait les mêmes intrants que le groupe B. Il s'agit ensuite d'estimer les résultats du groupe A si celui-ci avait les mêmes retours sur les intrants en termes de résultats que le groupe B. La méthode Blinder-Oaxaca permet donc de décomposer l'écart observé en une partie expliquée par des différences au niveau des variables explicatives et une partie expliquée par des différences au niveau du retour des variables en termes de variable expliquée. Cette partie est qualifiée d'écart inexplicé et comprend aussi l'effet des variables non-mesurées.

L'utilisation de la méthode Blinder-Oaxaca en économie de l'éducation est relativement récente (Ammermueller, 2007a ; McEwan et Marshall, 2004 ; Schneeweis, 2011).⁵ McEwan et Marshall (2004) ont utilisé les données de l'UNESCO sur les résultats à des tests standardisés d'élèves en Amérique latine afin de comparer le Mexique et Cuba. Ils ont trouvé que 31 % de l'écart en faveur de Cuba pouvait être expliqué par des différences au niveau des intrants, notamment l'effet des pairs et les intrants familiaux. Ammermueller (2007a) a utilisé les résultats des tests d'aptitudes PISA afin de comparer les immigrants et les natifs en Allemagne. Il a utilisé une extension de la méthode Blinder-Oaxaca afin d'effectuer la comparaison sur l'ensemble de la distribution.⁶ Les résultats de la décomposition ont montré que les natifs profitaient de meilleurs intrants familiaux que les immigrants et que l'écart expliqué était relativement plus important dans le bas de la distribution. L'effet des intrants expliquait entre 70 % de l'écart total au niveau du test d'aptitudes en mathématiques et 85 % de l'écart total au niveau du test d'aptitudes en sciences. Schneeweis (2011) a utilisé les données PISA et les données du TIMSS afin de comparer les résultats des immigrants et des natifs dans 160 pays. Les résultats ont montré des écarts statistiquement significatifs en faveur des natifs avec des écarts expliqués représentant plus de 90 % de l'écart observé dans certaines régions (notamment l'Europe du Nord) et moins de 10 % de l'écart observé dans d'autres régions (Afrique).

Relativement peu d'études en économie de l'éducation se sont penchées sur le cas du Canada. Corak et Lauzon (2009) ont étudié l'effet de la taille des classes et du temps d'éducation sur les résultats entre les immigrants et les natifs. Worswick (2004) a utilisé les données de *l'Enquête longitudinale nationale sur les enfants et les jeunes* aussi pour comparer les résultats des enfants d'immigrants aux enfants de parents natifs. Il a trouvé des écarts conditionnels statistiquement significatifs en faveur des enfants de parents natifs, mais ces écarts conditionnels tendaient à disparaître avec le temps chez les enfants d'immigrants qui parlaient déjà le français ou l'anglais lors de leur arrivée au Canada. Willms (2004) a utilisé les données des résultats PISA afin de décomposer la variance entre les écoles et à l'intérieur des écoles pour les provinces canadiennes.

En ce qui a trait aux minorités et majorités linguistiques du Canada, les études empiriques en économie se sont davantage concentrées sur les écarts de revenus (Grenier et Lacroix, 1986

⁵ Voir aussi Ammermueller (2007b) et Duque *et al.* (2012).

⁶ Voir Juhn *et al.* (1993) pour l'extension de la méthode Blinder-Oaxaca à la régression quantile.

; Lavoie et Saint-Germain, 1991 ; Shapiro et Stelcner 1997). Grenier et Lacroix (1986) ont trouvé des disparités de revenus en faveur des anglophones de la région d'Ottawa-Hull. Les auteurs ont souligné notamment que les anglophones unilingues de la région de la capitale nationale avaient en moyenne des revenus plus élevés que les francophones bilingues. Lavoie et Saint-Germain (1991) ont utilisé la langue parlée à la maison comme facteur discriminant et ont trouvé des écarts statistiquement significatifs au niveau des revenus totaux en faveur des anglophones dans l'ensemble du Canada, sauf au Québec où l'effet restait ambigu. Shapiro et Stelcner (1997) ont trouvé que l'écart salarial entre les anglophones et les francophones au Québec était disparu suite à l'adoption de la loi 101. Ils ont noté cependant que des écarts persistaient en défaveur des allophones et des francophones unilingues.

À la lumière de notre recension, une seule étude porte explicitement sur la comparaison des résultats sur des tests standardisés entre les élèves d'écoles de langue minoritaire et d'écoles de langue majoritaire du Canada. Shipley (2011) a utilisé les données des tests d'aptitudes PISA afin d'offrir un portrait descriptif des disparités entre les élèves d'écoles de langue minoritaire et d'écoles de langue majoritaire. Les résultats ont montré que les élèves d'écoles de langue majoritaire ont des résultats significativement supérieurs aux élèves d'écoles de langue minoritaire sur tous les tests d'aptitudes PISA, soit compréhension en lecture, raisonnement mathématique et sciences. Au niveau provincial, il n'y avait pas d'écart significatif au Québec entre les élèves d'écoles francophones et d'écoles anglophones sur l'ensemble des tests d'aptitudes. Les résultats des élèves d'écoles francophones étaient un demi écart-type inférieur à ceux des élèves d'écoles anglophones sur tous les tests d'aptitudes en Ontario, en Alberta et en Colombie-Britannique. Shipley (2011) conclut son étude en soulignant l'importance d'analyser davantage ces résultats afin de tenter de comprendre d'où provient l'écart observé entre les élèves d'écoles de langue minoritaire et d'écoles de langue majoritaire sur les tests d'aptitudes PISA. En effet, les résultats de Shipley (2011) ne nous permettent pas de voir de quoi se compose l'écart entre les deux groupes ni même de voir l'effet des caractéristiques propres aux deux groupes sur leurs résultats respectifs.

Le présent mémoire contribue à la littérature en offrant une analyse empirique de l'écart des résultats PISA entre les élèves d'écoles de langue minoritaire et d'écoles de langue majoritaire au Canada en utilisant la méthode Blinder-Oaxaca afin de décomposer l'écart entre les deux groupes. De plus, les résultats présentés dans ce mémoire constituent un outil important pour les

minorités linguistiques canadiennes en matière de défense des droits acquis en vertu de l'article 23 de la Charte canadienne des droits et libertés.

3. L'étude PISA

Les tests d'aptitudes en littératie PISA sont administrés tous les 3 ans depuis 2000 aux élèves de 15 ans dans les pays de l'OCDE et dans certains pays partenaires.⁷ Les tests visent à mesurer la compréhension et l'application de connaissances et d'aptitudes dans les domaines de la compréhension en lecture, du raisonnement mathématique et du raisonnement scientifique. Les tests sont d'une durée combinée de deux heures. Chaque cycle du test accorde une plus grande importance à un domaine en particulier. En 2009, la compréhension en lecture était le domaine approfondi, produisant ainsi des données plus précises et plus complètes sur les aptitudes des élèves dans ce domaine. Les tests d'aptitudes sont complétés par un questionnaire écrit de 30 minutes, répondu par les élèves, qui vise à rendre compte du statut socio-économique de l'élève ainsi que d'un questionnaire écrit de 30 minutes, répondu par les directeurs d'école, qui vise à rendre compte des spécifications de l'école fréquentée par l'élève. Bien que les tests d'aptitudes soient la pierre angulaire de l'étude PISA, les questionnaires administrés au niveau des élèves et au niveau de l'école sont fondamentaux dans la compréhension de l'ensemble des dimensions du processus éducatif.⁸

La méthode d'échantillonnage de l'étude PISA est de type stratifié et niché à deux niveaux. L'échantillon international est d'abord stratifié par pays et ensuite stratifié à l'intérieur des pays selon les demandes particulières de chaque pays. Au Canada, l'échantillon est stratifié au niveau des provinces et de la langue d'enseignement. Les écoles constituent le premier niveau de l'échantillon. Puisque seuls les élèves de 15 ans sont considérés dans l'étude, les écoles considérées lors de l'échantillonnage sont celles comprenant au moins un élève de 15 ans inscrit en 7e année ou plus (ou l'équivalent). Les écoles sont échantillonnées selon le nombre d'élèves inscrits dans l'école (échantillonnage probabiliste systématique proportionnel au nombre d'élèves inscrits dans l'école). Les élèves nichés dans les écoles constituent le deuxième niveau de

⁷ La liste des pays partenaires varie à chaque cycle de l'étude PISA. On note cependant une augmentation du nombre de pays partenaires depuis 2000 (15 pays partenaires en 2000 et 37 pays partenaires en 2009).

⁸ Un questionnaire destiné aux parents des élèves est aussi fourni par le consortium PISA, cependant le Canada n'a pas administré ce questionnaire lors du cycle de 2009. De plus, les pays sont invités à administrer des questionnaires optionnels, ce que le Canada a fait en 2009 avec un questionnaire visant à recueillir de l'information supplémentaire sur les élèves canadiens. Les données de ce questionnaire optionnel ne sont pas utilisées dans le présent mémoire.

l'échantillon. Les élèves sont échantillonnés avec les mêmes probabilités de sélection à partir d'une liste de l'ensemble des élèves de 15 ans inscrits dans les écoles échantillonnées au premier niveau.⁹ Le consortium PISA fournit un poids final pour chaque élève qui prend en considération la structure de l'échantillon stratifié d'élèves nichés dans les écoles. Le poids final est aussi ajusté pour prendre en considération le taux de non-réponse au niveau des écoles et au niveau des élèves. Conformément aux recommandations du consortium PISA (OCDE, 2009 : 56), les analyses comprises dans le présent mémoire ont été pondérées avec le poids final des élèves fournis dans la base de données publique.

Plusieurs éléments concernant la structure de l'échantillon et la nature des données doivent être pris en considération lors de l'analyse des données et de l'interprétation des résultats, notamment l'estimation de la variance, l'analyse des données des écoles au niveau des élèves et la nature des variables des résultats des tests d'aptitudes.

Premièrement, la structure de l'échantillon (élève niché dans l'école) pose un problème particulier dans l'estimation de la variance. En effet, les élèves d'une même école ne peuvent pas être considérés comme des observations indépendantes. Si l'objectif des analyses est d'inférer des résultats sur la population des élèves de 15 ans d'un pays particulier, il est primordial de prendre en considération le fait que la variance des élèves à l'intérieur d'une même école sera généralement plus faible que la variance entre les élèves d'écoles différentes. Afin de rendre compte de la variance dans l'école et de la variance entre les écoles dans les analyses, le consortium PISA recommande que l'estimation de la variance soit faite à partir de la méthode du *Balance Repeated Replication* (BRR).¹⁰ Cette méthode consiste à regrouper les écoles en paires pour ensuite accorder, de façon aléatoire, un poids particulier à chaque école dans chaque paire. Les deux poids utilisés dans chaque paire sont 0.5 et 1.5 suivant les recommandations de Judkins (1990) sur l'utilisation du correcteur de Fay.¹¹ Ce processus consiste donc en une réplification de l'échantillon avec un ajustement apporté à la probabilité de sélection de chaque école. Les poids finaux des élèves sont ensuite ajustés pour prendre en considération l'ajustement du poids des écoles. Le consortium PISA fournit 80 poids finaux ajustés, ou poids répliqués, pour chaque

⁹ Dans les pays comme le Canada, où les tests d'aptitudes ont été administrés durant les mois de mars, avril ou mai, les élèves qui avaient 15 ans au 1^{er} janvier 2009 ont été considérés dans le processus d'échantillonnage.

¹⁰ Voir Judkins (1990) pour une analyse plus approfondie de la méthode BRR pour l'estimation de la variance ainsi que du correcteur de Fay.

¹¹ Le correcteur de Fay permet notamment d'assurer la stabilité de l'estimation de la variance lorsque seul un sous-groupe de la population totale est considéré dans l'estimation (OCDE, 2012 : 126-127).

élève. L'estimation de la variance d'une statistique particulière implique donc de calculer cette statistique avec le poids final de chaque élève (statistique de référence), pour ensuite calculer la même statistique avec les 80 poids répliqués de chaque élève (statistiques répliquées). La variance estimée est ensuite calculée en prenant la différence entre les statistiques répliquées et la statistique de référence.¹² Cette méthode est convergente asymptotiquement et permet de rendre compte de la vraie variance de l'échantillon, c'est-à-dire une combinaison de la variance à l'intérieur de l'école et de la variance entre les écoles.

Le deuxième élément à prendre en considération repose sur le fait que l'échantillon n'a pas été construit dans l'optique de produire un portrait représentatif des écoles, mais bien un portrait représentatif des élèves de 15 ans qui fréquentent ces écoles. En effet, puisque les écoles ont été sélectionnées en fonction du nombre d'élèves de 15 ans inscrits dans l'école, certains types d'écoles vont être sous représentés dans l'échantillon si l'objectif est d'inférer des statistiques sur la population des écoles. Ainsi, le consortium PISA recommande de toujours interpréter les résultats d'analyses au niveau des élèves à défaut d'inférer des résultats biaisés au niveau des écoles. Cette recommandation est appliquée dans l'analyse des statistiques descriptives du présent mémoire.¹³

Le dernier élément à considérer concerne la nature des variables des résultats des tests d'aptitudes (OCDE, 2009 : 77-101). Afin de mesurer extensivement les connaissances et les aptitudes des élèves, plusieurs éléments de littératie doivent être évalués. Les évaluateurs font cependant face à des contraintes quant à la durée de la concentration des élèves et au temps alloué par les directeurs d'école pour administrer les tests. Les tests d'aptitudes ne peuvent donc pas comprendre l'ensemble des questions élaborées dans le cadre de l'étude. Le consortium PISA utilise donc des livrets de test aléatoires où seulement un sous-ensemble de questions est posé à chaque élève. Les résultats des tests sont ensuite mis à l'échelle suivant les spécifications de la théorie des réponses aux items et plus spécifiquement les spécifications du modèle de Rasch.¹⁴ Ce modèle permet de faire des imputations probabilistes sur les questions auxquelles l'élève n'a

¹² Formellement : $\hat{\sigma}^2_{(\hat{\theta})} = \frac{1}{G(1-k)^2} \sum_{i=1}^G (\hat{\theta}_{(i)} - \hat{\theta})^2$

où G est le nombre de répliqués et k est le correcteur de Fay (0.5 dans l'étude PISA).

¹³ Pour l'ensemble des analyses de régression, les variables du niveau école sont identifiées *de facto* comme des attributs des élèves.

¹⁴ Les spécifications du modèle de Rasch sont disponibles dans *PISA 2009 Technical Report* (OCDE, 2012 : 130-141).

pas répondu et de classer l'ensemble des élèves échantillonnés sur un continuum d'aptitudes. Une série de cinq valeurs plausibles pour chacun des domaines sont ensuite assignées à chaque élève. Ces valeurs plausibles sont des nombres tirés aléatoirement de la distribution marginale postérieure des résultats de l'élève.¹⁵ Les valeurs plausibles servent à rendre compte de la distribution du niveau d'aptitudes estimé de l'élève qui est une variable latente. L'utilisation des valeurs plausibles permet de produire des estimations non biaisées du niveau d'aptitudes des élèves. Puisque les valeurs plausibles sont des valeurs aléatoires, elles ont notamment leur propre variance (variance imputée). Le consortium PISA recommande d'utiliser l'ensemble des cinq valeurs plausibles dans l'estimation des statistiques relatives au niveau d'aptitudes et de combiner la variance imputée et la variance estimée dans l'analyse des résultats. Ainsi, pour chaque estimation du niveau d'aptitudes des élèves, le calcul doit être effectué sur chaque valeur plausible (5) et pour chaque réplification de l'échantillon (80). Cette procédure est appliquée dans l'ensemble des analyses du présent mémoire. Les valeurs plausibles rapportées dans la base de données PISA ont été standardisées afin d'avoir une moyenne internationale de 500 et un écart-type international de 100.

4. Données canadiennes et variables d'intérêts

L'échantillon canadien de l'étude PISA 2009 comprend 23 186 élèves répartis dans 977 écoles à travers les 10 provinces canadiennes.¹⁶ De ce nombre d'écoles, 827 écoles (comprenant 19 353 élèves) sont identifiées à des provinces particulières et peuvent donc être utilisées à des fins de comparaisons géographiques à l'intérieur du Canada. Puisque l'article 23 de la Charte ne concerne que les écoles publiques, nous ne considérons, dans notre analyse, que les écoles qui reçoivent du financement public pour mener leurs opérations, c'est-à-dire les écoles publiques et les écoles privées qui reçoivent du financement public. De plus, les provinces de Terre-Neuve et de l'Île-du-Prince-Édouard n'ont pas échantillonné d'écoles francophones et ne sont donc pas considérées dans notre analyse. Les groupes linguistiques sont identifiés par la langue utilisée

¹⁵ Voir Mislevy (1991) pour des explications plus approfondies sur les valeurs plausibles.

¹⁶ Une école a été retirée de l'échantillon original puisque qu'elle était la seule école à l'intérieur de sa strate ce qui rendait impossible l'estimation de la variance par BRR. Les territoires canadiens ne sont pas considérés dans l'étude PISA.

pour effectuer les tests d'aptitudes PISA, soit l'anglais ou le français.¹⁷ Nous parlons donc d'élèves d'écoles de langue minoritaire et d'élèves d'écoles de langue majoritaire.

Les fréquences et pourcentages de la répartition des écoles publiques et du nombre d'élèves par région et par groupe linguistique dans l'échantillon final sont présentés au Tableau 1. L'échantillon final comprend 16 041 élèves répartis dans 681 écoles. Il est intéressant de noter la surreprésentation des élèves d'écoles de langue minoritaire dans l'échantillon par rapport à la représentation des minorités linguistiques au niveau de la population. À titre indicatif, les données du recensement canadien de 2006 montrent que seulement 13.4 % des Québécois ont identifié l'anglais comme première langue officielle parlée alors que les élèves d'écoles anglophones représentent 23 % de l'échantillon final au Québec. Du côté des minorités francophones, le Nouveau-Brunswick était la seule province avec plus de 20 % des recensés qui disaient avoir appris le français comme première langue officielle, suivie par l'Ontario avec 4.5 % (Commissariat aux langues officielles, 2011) alors que les proportions d'élèves d'écoles francophones sont respectivement de 81 % et 35 % dans l'échantillon final. Cette surreprésentation des minorités linguistiques dans l'échantillon est le fruit de la structure de l'échantillonnage mis en place au Canada. Il en découle un niveau de précision accrue lors de l'estimation des statistiques sur les élèves qui fréquentent des écoles de langue minoritaire, ce qui améliore les comparaisons entre les groupes linguistiques et entre les provinces.

Plusieurs variables mesurées dans l'étude PISA sont utilisées dans le présent mémoire afin de procéder à des comparaisons entre les élèves d'écoles de langue minoritaire et d'écoles de langue majoritaire. À l'instar des études en économie de l'éducation précédemment exposées dans notre revue de littérature, les variables utilisées se situent à trois niveaux, soit le niveau de l'élève, le niveau familial et le niveau institutionnel. Les variables au niveau de l'élève et au niveau familial proviennent du questionnaire répondu par l'élève alors que les variables au niveau institutionnel proviennent du questionnaire répondu par le directeur de l'école (à l'exception de l'indice du niveau socio-économique moyen dans l'école qui est construit avec les réponses du questionnaire de l'élève). Les variables du niveau de l'élève sont les *résultats de tests de compréhension en lecture et de raisonnement mathématique, sexe, cheminement normal,*

¹⁷ Les systèmes d'éducation, francophone et anglophone, ont été considérés indépendamment dans l'échantillonnage des écoles. Les données publiques ne contiennent aucune information sur les programmes d'immersion en langue seconde. Nous notons cependant que la langue dans laquelle les tests d'aptitudes ont été administrés est la même pour tous les élèves d'une même école.

immigrant, et indice d'attitude envers l'école. Les variables du niveau familial sont *langue parlée à la maison, structure familiale, éducation de la mère, et nombre de livres à la maison.* Les variables du niveau institutionnel sont *population dans la communauté, ratio d'ordinateurs par élève dans l'école, ratio d'élèves par enseignant, nombre d'élèves dans l'école, indice du niveau socio-économique moyen dans l'école, et indice de l'attitude des enseignants.* Une description des variables et de la façon dont elles ont été traitées est présentée à l'Annexe 1.¹⁸

Les moyennes pondérées des variables utilisées dans le présent mémoire, réparties par régions géographiques et groupes linguistiques, sont présentées au Tableau 2. Nous pouvons voir que les élèves d'écoles de langue minoritaire ont des résultats significativement inférieurs à ceux des élèves d'écoles de langue majoritaire à l'exception des résultats de compréhension en lecture au Québec. Dans le cas des élèves d'écoles francophones hors Québec, le résultat moyen sur le test de compréhension en lecture est 50 points inférieurs à celui des élèves d'écoles anglophones, une différence d'un demi écart-type international. Malgré les résultats inférieurs des élèves d'écoles francophones hors Québec, l'indice moyen de l'attitude envers l'école des élèves francophones (0.23) est supérieur à celui des élèves anglophones (0.06). Notons aussi la disparité entre les élèves d'écoles de langue minoritaire et d'écoles de langue majoritaire en ce qui a trait à la langue parlée à la maison. Les élèves d'écoles de langue minoritaire parlent moins la langue du test à la maison. Dans le Canada hors Québec, seulement 57 % des élèves d'écoles francophones minoritaires parlent le français à la maison alors que 86 % des élèves d'écoles anglophones majoritaires parlent l'anglais à la maison. Du côté du Québec, la situation est similaire, quoique l'écart soit moins important, alors que 76 % des élèves d'écoles anglophones minoritaires parlent l'anglais à la maison et 90 % des élèves d'écoles francophones majoritaires parlent le français à la maison. Les élèves d'écoles de langue minoritaire fréquentent également des écoles en moyenne plus petites que les élèves d'écoles de langue majoritaire.

Les données montrent également des disparités entre les régions en ce qui a trait aux caractéristiques des élèves d'écoles de langue minoritaire relativement aux élèves d'écoles de langue majoritaire. Les élèves d'écoles francophones du Canada hors Québec habitent dans des communautés moins peuplées que les élèves d'écoles anglophones. Ils fréquentent des écoles où le niveau socio-économique moyen est moins élevé que les élèves d'écoles anglophones. Ils

¹⁸ Aucune des variables utilisées ne comprend des valeurs manquantes totalisant plus de 7.63 % des observations totales.

ont en moyenne moins de livres à la maison que les élèves d'écoles anglophones. Ils accusent également, en moyenne, un retard sur le cheminement normal supérieur à celui des élèves qui fréquentent une école anglophone. Au Québec, la situation est inversée pour l'ensemble de ces variables. En moyenne, les élèves d'écoles anglophones habitent dans des communautés plus peuplées, fréquentent des écoles où le niveau socio-économique moyen est plus élevé, ont plus de livres à la maison et accusent un retard sur le cheminement normal inférieur à celui des élèves qui fréquentent une école francophone.

5. Méthode empirique

Afin d'analyser l'effet des caractéristiques de chaque groupe sur leurs résultats respectifs, nous construisons une fonction de production d'éducation de la forme :

$$Y_{is} = \beta_0 + \beta_1 E_{is} + \beta_2 F_{is} + \beta_3 I_s + \nu_s + \varepsilon_{is}$$

où Y_{is} est le résultat de l'étudiant (i) dans l'école (s), E_{is} est un vecteur des variables au niveau de l'élève, F_{is} est un vecteur de variables au niveau familial, I_s est un vecteur de variables au niveau institutionnel, ν_s est le terme d'erreur au niveau de l'école et ε_{is} est le terme d'erreur au niveau de l'élève.¹⁹ La fonction de production d'éducation nous permet d'identifier les caractéristiques qui expliquent les résultats. Notons que, puisque les élèves sont nichés dans les écoles, les caractéristiques au niveau de l'élève et les caractéristiques au niveau familial ne sont pas indépendantes du terme d'erreur au niveau de l'école (ν_s). Afin de contourner ce biais potentiel, nous estimons la variance des paramètres de la régression avec la méthode *BRR* en utilisant les poids répliqués fournis dans la base de données publique par le consortium PISA. Cette méthode nous permet d'obtenir des erreurs standards qui sont robustes à la structure de l'échantillon (OCDE, 2009 : 65-74).²⁰ L'estimation des paramètres est aussi pondérée par le poids échantillonnel final de chaque élève. La fonction de production d'éducation est estimée séparément pour les élèves d'écoles de langue minoritaire et les élèves d'écoles de langue majoritaire.

¹⁹ Cette forme de fonction de production d'éducation est amplement utilisée dans la littérature en économie de l'éducation (Ammermueller, 2007a ; Hanushek et Woessmann, 2010 ; Schneeweis, 2011).

²⁰ Bien que la méthode *BRR* nous permette de consolider les termes d'erreur au niveau de l'élève, cela ne veut pas dire pour autant qu'elle corrige entièrement les problèmes d'endogénéité du modèle.

Afin de comparer les fonctions de production d'éducation estimées pour les deux groupes, nous utilisons la méthode de décomposition Blinder-Oaxaca (Blinder, 1973 ; Oaxaca, 1973). Cette méthode de décomposition a originalement été élaborée afin d'expliquer l'écart salarial entre les hommes et les femmes. Depuis les années 1990, cette méthode de décomposition a amplement été utilisée dans le contexte des fonctions de production d'éducation. L'idée derrière la décomposition Blinder-Oaxaca est de déterminer quelle proportion de l'écart observé entre les deux groupes au niveau de la variable prédite peut être expliquée par des différences au niveau des caractéristiques de l'élève, des caractéristiques familiales, et des caractéristiques institutionnelles et quelle proportion de l'écart reste inexpliquée.

Notons par $\bar{Y}_{maj} = \bar{X}'_{maj} \hat{\beta}_{maj}$ (respectivement $\bar{Y}_{min} = \bar{X}'_{min} \hat{\beta}_{min}$) l'estimation de la fonction de production d'éducation des élèves des écoles de la majorité (minorité) linguistique évaluée à la moyenne où \bar{Y}_{maj} (\bar{Y}_{min}) est la valeur moyenne du résultat sur le test d'aptitudes, $\hat{\beta}_{maj}$ ($\hat{\beta}_{min}$) est un vecteur de coefficients estimés et \bar{X}_{maj} (\bar{X}_{min}) est un vecteur comprenant l'ensemble des variables du modèle évaluées à la moyenne.²¹ L'écart entre les élèves des écoles de la minorité linguistique et les élèves des écoles de la majorité linguistique peut être exprimé comme suit ;

$$\Delta \bar{Y}_{maj-min} = \bar{X}'_{maj} \hat{\beta}_{maj} - \bar{X}'_{min} \hat{\beta}_{min} = \hat{\beta}_{maj} (\bar{X}_{maj} - \bar{X}_{min})' + \bar{X}'_{min} (\hat{\beta}_{maj} - \hat{\beta}_{min})$$

où $\hat{\beta}_{maj} (\bar{X}_{maj} - \bar{X}_{min})'$ représente l'écart expliqué par une différence au niveau des caractéristiques de l'élève (écart expliqué), et $\bar{X}'_{min} (\hat{\beta}_{maj} - \hat{\beta}_{min})$ représente l'écart dû à une différence au niveau du retour des caractéristiques en termes de résultats (écart inexpliqué).

L'écart inexpliqué, qui inclut la constante du modèle, capture donc la différence au niveau des caractéristiques inobservées. En économie du travail, le terme « discrimination » a souvent été utilisé pour caractériser l'écart inexpliqué.²² Or, en économie de l'éducation, on ne peut proprement parler de discrimination envers un groupe. L'interprétation de l'écart inexpliqué dans le cadre du présent mémoire repose sur l'idée qu'un groupe peut faire face à des difficultés

²¹ Par construction, $\overline{(v_s + \varepsilon_{is})} = 0$.

²² Le terme discrimination est aussi critiqué en économie du travail du fait que l'écart inexpliqué capture l'effet des variables qui ne sont pas incluses dans le modèle (Jann, 2008).

particulières qui ne sont pas capturées par les données observées. Lorsque l'écart des résultats estimés est décomposé, les difficultés particulières sont capturées dans l'écart inexpliqué.²³

Un problème souvent rencontré avec la décomposition Blinder-Oaxaca repose sur le choix du groupe de référence. En effet, dans certaines situations, il peut être ardu de déterminer si un des groupes est désavantagé par rapport à l'autre ou si un des groupes est favorisé par rapport à l'autre groupe. Par exemple, il convient de se demander si l'écart salarial entre les hommes et les femmes est dû à une discrimination négative envers les femmes par rapport aux hommes ou à une discrimination positive envers les hommes par rapport aux femmes. Dans le cadre de l'économie de l'éducation, l'utilisation d'un groupe de référence repose sur le fait que le niveau de résultats moyens le plus élevé représente le niveau de résultats moyens de référence. L'écart est ainsi mesuré seulement en termes de désavantage d'un groupe par rapport à l'autre groupe.

Ainsi l'écart expliqué, dans le présent mémoire, peut être interprété comme l'écart qui persisterait entre les élèves des écoles de la minorité linguistique et des écoles de la majorité linguistique si les élèves des écoles de la minorité linguistique avaient les mêmes retours sur les caractéristiques en termes de résultats que les élèves des écoles de la majorité linguistique. Il en résulte donc un écart entre les deux groupes simplement dû à des différences au niveau des variables explicatives. Par ailleurs, l'écart inexpliqué capture l'écart qui persisterait si les élèves des écoles de la majorité linguistique avaient les mêmes caractéristiques moyennes que les élèves des écoles de la minorité linguistique. Il en résulte donc un écart simplement dû à des différences au niveau des retours sur les caractéristiques en termes de résultats et des différences au niveau de la constante du modèle. Il est important de noter que les écarts, expliqués et inexpliqués, peuvent être positifs ou négatifs. Dans le cas d'un écart positif, nous dirons qu'il y a augmentation de l'écart total observé entre les deux groupes alors que pour un écart négatif nous dirons qu'il y a diminution de l'écart total observé entre les deux groupes.

Des fonctions de production d'éducation ont été estimées pour les élèves des écoles de la minorité linguistique et des écoles de la majorité linguistique du Canada hors Québec, du Québec, de l'Ontario et de l'Alberta et de la Colombie-Britannique.^{24 25} L'écart entre les élèves

²³ Schneeweis (2009) utilise une interprétation similaire de l'écart inexpliqué pour analyser l'écart entre les immigrants et les natifs au niveau des résultats sur des tests d'aptitudes.

²⁴ Les analyses ont été effectuées avec le logiciel statistique *Stata 12.1*.

²⁵ Nous utilisons le module *-pv-* afin de considérer les valeurs plausibles dans l'estimation des fonctions de production d'éducation. Le module *-pv-* est un programme écrit par Macdonald (2008) dans le but de procéder à des

des écoles de langue minoritaire et des écoles de langue majoritaire dans chacune des régions a ensuite été décomposé selon la méthode Blinder-Oaxaca. Les résultats des analyses sont présentés et discutés dans la prochaine section.²⁶

6. Résultats

6.1 Les fonctions de production d'éducation

6.1.1 Compréhension en lecture

Les résultats de l'estimation de la fonction de production d'éducation pour le test de compréhension en lecture pour les élèves d'écoles de langue minoritaire et d'écoles de langue majoritaire du Canada hors Québec et du Québec sont présentés au Tableau 3. Il est important de noter que les coefficients représentent ici des mesures d'association et non des effets causals.²⁷ Les résultats pour les élèves des écoles de langue minoritaire et des écoles de langue majoritaire des deux régions géographiques montrent un effet statistiquement significatif de plusieurs variables.²⁸ La variable *sexe* est associée à une différence statistiquement significative des résultats moyens de compréhension en lecture entre les filles et les garçons. C'est donc dire que les filles ont en moyenne des résultats de 20.18 (élèves d'écoles francophones du Québec) à 27.73 (élèves d'écoles francophones du Canada hors Québec) supérieurs à ceux des garçons. On note aussi peu de variation au niveau de la variable *nombres de livres à la maison* à travers les groupes et à travers les régions géographiques. Le fait d'avoir déclaré avoir plus de 100 livres à la maison est associé à une augmentation du résultat moyen allant de 32.27 points pour les élèves d'écoles francophones du Québec à 38.69 points pour les élèves d'écoles anglophones du Canada hors Québec. Pour ce qui est des variables *cheminement normal* et *niveau socio-économique moyen* dans l'école, les résultats sont beaucoup plus variables entre les groupes et les régions

analyses en considérant les valeurs plausibles. Ce module a été conçu spécifiquement pour analyser les données PISA.

²⁶ Le module -oaxaca- est utilisé afin de décomposer l'écart entre les élèves des écoles de langue minoritaire et des écoles de langue majoritaire au niveau des résultats sur les tests d'aptitudes. Ce module est un programme écrit par Jann (2008) dans le but d'appliquer la décomposition Blinder-Oaxaca aux analyses effectuées avec *Stata*. Le module -oaxaca- utilise la « méthode delta » afin de produire une approximation de l'erreur standard des écarts estimés qui soit convergente asymptotiquement (Jann, 2008). L'utilisation du module -oaxaca- nous permet donc de faire des inférences sur la population en ce qui a trait aux écarts estimés. Duque *et al.* (2009) utilisent notamment ce module pour décomposer l'écart au niveau des résultats PISA entre les élèves urbains et les élèves ruraux en Colombie.

²⁷ Hanushek et Woessmann (2010) soulignent la nature approximative des variables mesurées dans le cadre des études sur les déterminants de l'éducation. Les variables approximatives (*proxy*) incluses dans la fonction de production visent à rendre compte de facteurs causals inobservés. Or, même si une variable approximative est présumée fortement associée à un facteur causal, l'effet de la variable approximative n'est pas causal pour autant.

²⁸ L'effet estimé des variables est interprété *ceteris paribus*.

géographiques. Une variation d'un écart-type du niveau socio-économique moyen dans l'école est associée à une augmentation du résultat moyen allant de 25.66 points pour les élèves d'écoles francophones du Canada hors Québec à 46.63 points pour les élèves d'écoles anglophones du Québec. Le fait pour un élève d'être à temps sur le cheminement normal est associé à une augmentation du résultat moyen allant de 35.45 points pour les élèves d'écoles anglophones du Québec à 59.94 points pour les élèves d'écoles francophones du Canada hors Québec. Les variables *immigrant* et *ratio d'ordinateurs par élève* n'ont aucun effet statistiquement significatif dans l'ensemble des groupes et des régions géographiques.

Les résultats montrent aussi des différences au niveau des déterminants des résultats entre les groupes et entre les régions géographiques. Au niveau de la variable de *l'indice d'attitude envers l'école*, l'association entre la variable et les résultats est significative à 5 % pour les élèves d'écoles francophones du Québec et du Canada hors Québec ainsi que pour les élèves d'écoles anglophones du Canada hors Québec, mais est faiblement significative (10 %) pour les élèves d'écoles anglophones du Québec. La *langue parlée à la maison* est aussi associée à un résultat moyen plus élevé pour les élèves qui parlent la langue du test à la maison pour l'ensemble des groupes à l'exception des élèves d'écoles anglophones du Canada hors Québec. Au niveau de la *structure familiale* et de *l'éducation de la mère*, les résultats sont positifs et statistiquement significatifs pour les élèves du Canada hors Québec, mais ne sont pas statistiquement significatifs pour les élèves du Québec. L'effet de la variable de *population dans la communauté* est significatif à 5 % et positif pour les élèves du Canada hors Québec et faiblement significatif (10 %) pour les élèves d'écoles anglophones du Québec sans être statistiquement significatif pour les élèves d'écoles francophones de la même province. Un ratio de 17 élèves ou plus par enseignant est associé à des résultats moyens plus élevés, et statistiquement significatif, pour les élèves d'écoles de langue minoritaire, quoique faiblement significatif (10 %) pour les élèves d'écoles francophones du Canada hors Québec. Finalement, le fait de fréquenter une école de 1000 élèves et plus (*nombre d'élèves dans l'école*) est associé à des résultats moyens plus élevés et statistiquement significatifs pour les élèves d'écoles francophones du Canada hors Québec, mais moins élevés pour les élèves d'écoles anglophones du Québec.

6.1.2 Raisonnement mathématique

Les résultats de l'estimation de la fonction de production d'éducation pour le test de raisonnement mathématique pour les minorités et majorités linguistiques du Canada hors Québec et du Québec sont présentés au Tableau 4. Tout comme dans le modèle précédent, l'effet des variables de *sexe*, du *cheminement normal*, du *nombre de livres à la maison* ainsi que du *niveau socio-économique moyen dans l'école* est statistiquement significatif pour l'ensemble des groupes des deux régions géographiques. Notons cependant que l'effet de la variable *sexe*, qui était positif dans le modèle précédent, est négatif dans le présent modèle. C'est donc dire que les filles ont en moyenne des résultats de 17.12 (élèves d'écoles anglophones du Canada hors Québec) à 27.82 (élèves d'écoles francophones du Québec) inférieurs à ceux des garçons. Les résultats en ce qui a trait aux variables *immigrant* et *ratio d'ordinateurs par élève* sont toujours non concluants pour l'ensemble des groupes linguistiques. S'ajoutent à celles-ci les variables *ratio d'élèves par enseignant* et *langue parlée à la maison*, la dernière étant cependant faiblement significative (10 %) dans le cas des élèves d'écoles francophones du Québec.

Encore une fois, les résultats du présent modèle montrent des différences entre les groupes linguistiques. *L'indice d'attitude envers l'école* n'est statistiquement significatif que pour les élèves de langue majoritaire. La taille d'effet est relativement faible sur les résultats dans les deux cas. Une variation positive d'un écart-type de l'indice est associée à une augmentation des résultats moyens de 6.74 et 5.49 points pour les élèves d'écoles anglophones du Canada hors Québec et les élèves d'écoles francophones du Québec respectivement. À l'instar des résultats du modèle précédent, l'effet des variables *structure familiale* et *éducation de la mère* est positif et fortement statistiquement significatif pour les groupes linguistiques du Canada hors Québec. On note aussi que l'effet de la variable *nombre d'élèves dans l'école* n'est plus statistiquement significatif pour les élèves d'écoles francophones du Canada hors Québec en ce qui a trait aux résultats en raisonnement mathématique. L'effet de la variable de *population dans la communauté* n'est aussi plus statistiquement significatif pour les élèves du Canada hors Québec en ce qui a trait aux résultats de raisonnement mathématique. Finalement, la variable de *l'indice de l'attitude des enseignants* est statistiquement significative pour les élèves d'écoles anglophones du Québec alors qu'elle n'est pas statistiquement significative pour les élèves des autres groupes linguistiques, et ce pour les deux tests confondus.

En somme, les caractéristiques individuelles semblent expliquer une partie substantielle de la variation dans les résultats, suivies par les caractéristiques familiales et les caractéristiques institutionnelles. Les filles réussissent en moyenne mieux que les garçons sur le test d'aptitudes de compréhension en lecture et moins bien que les garçons en raisonnement mathématique. Cette constatation est une régularité empirique dans la littérature sur les déterminants des résultats au niveau secondaire (Woessmann, 2003 ; Fuchs et Woessmann, 2007). Au plan du cheminement de l'élève, les élèves qui accusent un retard sur les autres élèves du même âge en ce qui a trait au niveau scolaire réussissent en moyenne moins bien que les élèves qui n'accusent pas de retard ou qui sont en avance. Dans la majorité des cas, l'écart en termes de résultats entre les deux types d'élèves avoisine le demi écart-type mondial. Ce type de résultats est aussi une régularité empirique dans la littérature (Ammermueller, 2007b ; Fuchs et Woessmann, 2007). Les résultats pour la variable *immigrant*, c'est-à-dire les élèves qui sont nés à l'extérieur du Canada par opposition aux élèves nés au Canada, sont particulièrement intéressants. Nos résultats montrent qu'il n'y a pas de différence significative à 5 % entre les élèves nés à l'extérieur du pays et les élèves nés au Canada en termes de résultats sur les tests d'aptitudes de compréhension en lecture et de raisonnement mathématique, et ce pour l'ensemble des groupes linguistiques et des régions géographiques. Ces résultats sont cohérents avec les résultats de Worswick (2004) qui trouve que les résultats en lecture et en mathématiques sont similaires chez les élèves de parents immigrants et les élèves de parents natifs du Canada.

Au niveau des caractéristiques familiales, seul l'effet du nombre de livres à la maison est régulier entre les groupes linguistiques, régions géographiques et tests d'aptitudes. Le nombre de livres à la maison est une variable approximative du niveau socioculturel de la famille qui vise à rendre compte des ressources éducatives disponibles à la maison et de l'importance de l'éducation et de la littératie pour les parents. Comme le soulignent Hanushek et Woessmann (2010), cette variable est un déterminant important du résultat des élèves dans la majorité des pays du monde et est par conséquent une variable importante dans la comparaison entre les pays. Nos résultats montrent que les élèves qui déclarent avoir plus de 100 livres à la maison ont en moyenne des résultats de plus d'un quart d'écart-type mondial supérieurs aux élèves qui déclarent avoir 100 livres et moins à la maison. Quant à la langue parlée à la maison, les résultats d'estimation sont très différents selon que l'on considère les résultats sur le test de lecture ou de mathématiques. Dans le cas des résultats pour la compréhension en lecture, le fait de parler la

langue du test à la maison a un effet positif et significatif à 5 % pour les élèves d'écoles francophones des deux régions géographiques et pour les élèves d'écoles anglophones du Québec. Dans le cas des résultats sur le test de raisonnement mathématique, l'effet n'est pas significatif à 5 % pour l'ensemble des groupes linguistiques.

Au niveau des caractéristiques institutionnelles, seul l'indice du niveau socio-économique moyen dans l'école est régulier entre les groupes linguistiques. Dans tous les cas, les élèves qui fréquentent une école où l'indice moyen du niveau socio-économique est d'un écart-type international supérieur à la moyenne internationale ont en moyenne des résultats plus élevés de 0.25 à 0.5 écart-type international par rapport aux élèves qui fréquentent une école où l'indice moyen du niveau socio-économique est la moyenne internationale. Les résultats suggèrent qu'il y a présence d'un effet de pairs dans les écoles publiques canadiennes, après avoir contrôlé pour les tierces variables. Les élèves qui fréquentent des écoles où le niveau socio-économique moyen est plus élevé ont aussi, en moyenne, des résultats plus élevés. Notons aussi l'effet négatif de la variable du nombre d'élèves dans l'école sur les résultats des élèves des minorités linguistiques. Les élèves d'écoles anglophones du Québec qui fréquentent une école de plus de 1000 élèves ont en moyenne des résultats significativement moins élevés que ceux qui fréquentent une école de 1000 élèves ou moins. D'un autre côté, les élèves d'écoles francophones du Canada hors Québec qui fréquentent une école de plus de 1000 élèves ont en moyenne des résultats en lecture significativement plus élevés que ceux qui fréquentent des écoles de 1000 élèves ou moins. Cette asymétrie en ce qui a trait à l'effet du nombre d'élèves dans l'école entre les élèves d'écoles de langue minoritaire du Canada hors Québec et du Québec témoigne d'une différence potentielle entre les systèmes d'éducation de langue minoritaire au Canada.

6.2 Les décompositions Blinder-Oaxaca

6.2.1 Compréhension en lecture

Les résultats de la décomposition Blinder-Oaxaca pour le test de compréhension en lecture sont présentés au Tableau 5. Pour le Canada hors Québec, l'écart estimé entre les élèves d'écoles francophones et les élèves d'écoles anglophones est statistiquement significatif et de 50.16 points, soit environ un demi écart-type international. Environ 17 % de l'écart (8.48 points) est expliqué par une différence au niveau des caractéristiques incluses dans le modèle alors que 83 % (41.68 points) de l'écart demeure inexpliqué. La décomposition de l'écart expliqué nous

montre que les variables de *sexe* et d'*attitude envers l'école* favorisent les élèves d'écoles francophones minoritaires. C'est donc dire que, si les élèves d'écoles francophones minoritaires avaient les mêmes compositions moyennes sur ces deux variables que les élèves d'écoles anglophones, ils obtiendraient, en moyenne, des résultats inférieurs. Les effets combinés des variables *livres à la maison*, *population dans la communauté* et *niveau socio-économique moyen dans l'école* contribuent à exacerber l'écart estimé. C'est donc dire que, si les élèves d'écoles francophones minoritaires avaient les mêmes compositions moyennes que les élèves de la majorité anglophone sur ces trois variables, ils obtiendraient, en moyenne, des résultats supérieurs. Dans le cas de la variable *livres à la maison*, si la composition moyenne des élèves qui possèdent plus de 100 livres à la maison était la même chez les élèves d'écoles francophones minoritaires que chez les élèves d'écoles anglophones majoritaires, l'écart estimé diminuerait de 4.39 points. Il demeure cependant que, dans le cas du Canada hors Québec, l'écart expliqué est relativement faible par rapport à l'écart inexpliqué.

Pour le Québec, les résultats montrent que l'écart n'est pas statistiquement significatif entre les élèves d'écoles anglophones minoritaires et les élèves d'écoles francophones majoritaires. Cependant, la décomposition de l'écart montre un écart expliqué négatif et statistiquement significatif ainsi qu'un écart inexpliqué positif et statistiquement significatif. C'est donc dire que, si les élèves d'écoles anglophones avaient les mêmes caractéristiques moyennes que les élèves d'écoles francophones (écart expliqué), ils obtiendraient en moyenne des résultats 14.97 points inférieurs aux résultats moyens observés. Les variables *cheminement normal*, *nombre de livres à la maison* et *indice du niveau socio-économique moyen* composent la plus grande part de l'écart expliqué. Notamment, si les élèves d'écoles anglophones avaient en moyenne le même niveau socio-économique moyen dans l'école que les élèves d'écoles francophones, l'écart estimé entre les deux groupes augmenterait de 7.58 points. Or, malgré le fait que l'écart expliqué avantage les élèves d'écoles anglophones, cet avantage est entièrement neutralisé par l'écart inexpliqué. L'effet des différences au niveau des caractéristiques observées est donc entièrement compensé par des différences quant au retour sur les caractéristiques en termes de résultats et par des différences sur les caractéristiques inobservées.

6.2.2 Raisonnement mathématique

Les résultats de la décomposition Blinder-Oaxaca pour le test d'aptitudes en raisonnement mathématique sont présentés au Tableau 6. Pour le Canada hors Québec, l'écart estimé entre les élèves d'écoles francophones et les élèves d'écoles anglophones est statistiquement significatif et de 23.31 points, soit environ un quart d'écart-type international. Environ 25 % de l'écart (5.88 points) est expliqué par une différence au niveau des caractéristiques incluses dans le modèle alors que 75 % (17.43 points) de l'écart demeure inexpliqué. La décomposition de l'écart expliqué nous montre que seule la variable d'*attitude envers l'école* favorise les élèves d'écoles francophones minoritaires. C'est donc dire que, si les élèves d'écoles francophones avaient en moyenne le même indice d'attitude envers l'école que les élèves d'écoles anglophones, ils obtiendraient, en moyenne, des résultats inférieurs. Cependant, les effets combinés des variables *cheminement normal*, *livres à la maison*, et *niveau socio-économique moyen dans l'école* contribuent à faire augmenter l'écart estimé. C'est donc dire que, si les élèves d'écoles francophones minoritaires avaient les mêmes compositions moyennes que les élèves de la majorité anglophone sur ces trois variables, ils obtiendraient, en moyenne, des résultats supérieurs. Dans le cas de la variable *livres à la maison*, si la composition moyenne des élèves qui possèdent plus de 100 livres à la maison était la même chez les élèves d'écoles francophones minoritaires que chez les élèves d'écoles anglophones majoritaires, l'écart estimé diminuerait de 4.42 points. Il demeure cependant que, dans le cas du Canada hors Québec, l'écart expliqué demeure relativement faible par rapport à l'écart inexpliqué.

Pour le Québec, les résultats montrent un écart estimé statistiquement significatif de 13.5 points en faveur des élèves d'écoles francophones majoritaires. La décomposition de l'écart montre cependant que l'écart expliqué est négatif et statistiquement significatif et que l'écart inexpliqué est positif et statistiquement significatif. C'est donc dire que, si les élèves d'écoles anglophones avaient les mêmes caractéristiques moyennes que les élèves d'écoles francophones (écart expliqué), ils obtiendraient en moyenne des résultats 17.1 points inférieurs au résultat observé. Les variables *cheminement normal*, *nombre de livres à la maison* et *indice du niveau socio-économique moyen* composent la plus grande part de l'écart expliqué. Notamment, si les élèves d'écoles anglophones avaient en moyenne le même niveau socio-économique moyen dans l'école que les élèves d'écoles francophones, l'écart estimé entre les deux groupes augmenterait de 9.77 points. Or, malgré le fait que l'écart expliqué avantage les élèves d'écoles anglophones,

cet avantage est entièrement neutralisé par l'écart inexpliqué, au point de créer un écart total statistiquement significatif en faveur des élèves d'écoles francophones. L'effet des différences au niveau des caractéristiques observées est donc entièrement compensé par des différences quant aux retours sur les caractéristiques en termes de résultats et par des différences au niveau des caractéristiques inobservées.

6.3 Les décompositions Blinder-Oaxaca pour l'Ontario, l'Alberta et la Colombie-Britannique

La décomposition Blinder-Oaxaca a aussi été faite pour les provinces de l'Ontario, de l'Alberta et de la Colombie-Britannique.²⁹ Dans les trois provinces, les élèves d'écoles francophones sont en situation minoritaire par rapport aux élèves d'écoles anglophones, néanmoins les provinces diffèrent quant au nombre d'élèves d'écoles francophones minoritaires, plus important en Ontario, et à la concentration géographique des francophones minoritaires, plus élevée en Ontario. Les résultats de la décomposition Blinder-Oaxaca pour le test d'aptitudes de compréhension en lecture pour les provinces de l'Ontario, de l'Alberta et de la Colombie-Britannique sont présentés au Tableau 7. L'écart estimé entre les élèves d'écoles francophones minoritaires et les élèves d'écoles anglophones majoritaires est positif et statistiquement significatif pour l'ensemble des trois provinces. C'est donc dire qu'en moyenne les résultats des élèves d'écoles francophones sont inférieurs à ceux des élèves d'écoles anglophones. L'analyse de décomposition montre que l'écart expliqué n'est pas statistiquement significatif pour les trois provinces. L'écart inexpliqué est cependant positif et statistiquement significatif, représentant 84.5 % de l'écart total en Ontario, 88 % de l'écart total en Alberta et presque la totalité de l'écart total en Colombie-Britannique ce qui montre que la proportion relative de l'écart inexpliqué par rapport à l'écart total est plus importante dans les provinces où les francophones minoritaires sont moins nombreux et moins concentrés géographiquement (Alberta et Colombie-Britannique).³⁰

7. Discussion

L'objectif de ce mémoire était de tester la présence d'un effet minoritaire chez les élèves qui fréquentent des écoles francophones du Canada hors Québec et chez les élèves qui fréquentent des écoles anglophones du Québec par rapport aux élèves qui fréquentent des écoles

²⁹ Les autres provinces ont été omises du fait que la taille de l'échantillon y est trop faible.

³⁰ La décomposition au niveau des provinces pour ce qui est des résultats en raisonnement mathématique n'est pas présentée puisque les résultats de la décomposition sont similaires à ceux de la compréhension lecture.

de la langue majoritaire respective. La présence d'un tel effet a été testée en analysant les résultats de tests d'aptitudes de compréhension en lecture et de raisonnement mathématique administrés dans le cadre de l'étude PISA en 2009. Avant de résumer les résultats, il convient d'apporter quelques précisions sur cet effet minoritaire.

L'effet minoritaire attendu qui comprend l'ensemble des difficultés particulières aux élèves dont la langue d'éducation est la langue de la minorité, est analogue au concept de mesure d'intégration de Schneeweis (2011). Schneeweis (2011) utilise la mesure d'intégration afin de caractériser l'écart inexplicé dans la décomposition des résultats sur des tests d'aptitudes entre immigrants et natifs. La situation des immigrants est analogue à celle des élèves qui fréquentent des écoles de langue minoritaire dans la mesure où les immigrants doivent aussi faire face à des difficultés particulières qui leur sont propres, notamment l'apprentissage d'une nouvelle culture et, dans certains cas, l'apprentissage d'une nouvelle langue. Cependant, les élèves qui fréquentent des écoles de langue minoritaire au Canada se distinguent du fait qu'ils ont accès à des institutions adaptées à leur situation minoritaire et que ces institutions sont garanties par la Charte des droits et libertés. Les élèves qui fréquentent des écoles de langue minoritaire ne font donc pas face à des difficultés d'intégration, mais bien à des difficultés qui découlent du fait qu'ils évoluent quotidiennement dans un contexte social où la langue d'éducation n'est pas la langue parlée par la majorité démographique.

Les résultats présentés dans la section précédente montrent que les élèves qui fréquentent une école de langue minoritaire réussissent en moyenne significativement moins bien que les élèves qui fréquentent une école de langue majoritaire, à l'exception des résultats de compréhension en lecture pour les élèves qui fréquentent des écoles anglophones au Québec. Il existe donc des écarts statistiquement significatifs en défaveur des élèves qui fréquentent des écoles de langue minoritaire sur les résultats de tests d'aptitudes PISA en lecture (sauf au Québec) et en mathématiques. La décomposition de l'écart entre les groupes (minoritaire ou majoritaire) et les régions géographiques (Canada hors Québec et Québec) montrent qu'une partie de l'écart n'est pas expliquée par des différences au niveau des caractéristiques individuelles, familiales ou institutionnelles incluses dans la fonction de production d'éducation. Cet écart inexplicé défavorise tous les élèves des écoles de langue minoritaire, soit des écoles francophones du Canada hors Québec et des écoles anglophones du Québec, par rapport aux élèves des écoles de la langue majoritaire respective. Dans le cas des élèves d'écoles anglophones

du Québec, l'écart inexpliqué compense entièrement l'écart expliqué dans les deux décompositions Blinder-Oaxaca (compréhension en lecture et raisonnement mathématique). Ce qui veut dire que, toutes choses étant égales par ailleurs, les avantages liés aux caractéristiques des élèves d'écoles anglophones en termes de résultats sont neutralisés par les désavantages liés aux retours des caractéristiques en termes de résultats et par des caractéristiques inobservées.

L'écart total en lecture est plus important que l'écart total en mathématiques pour les élèves d'écoles francophones du Canada hors Québec alors que cette situation est inversée pour les résultats des élèves d'écoles anglophones du Québec. De plus, la proportion de l'écart qui est inexplicée chez les élèves d'écoles francophones minoritaires est relativement plus importante pour les résultats de compréhension en lecture que pour les résultats de raisonnement mathématique. C'est dire que les matières scolaires qui portent plus spécifiquement sur la manipulation de la langue sont plus touchées par l'effet minoritaire chez les élèves d'écoles francophones au Canada hors Québec seulement. Ces résultats ne sont pas observés chez les élèves d'écoles anglophones du Québec. Ceci pourrait s'expliquer en partie par le fait que les élèves d'écoles anglophones du Québec profitent d'une situation particulière au plan linguistique en Amérique du Nord. En effet, même s'ils sont en situation minoritaire au niveau provincial, ils profitent notamment d'un bassin nord-américain majoritairement anglophone. L'accès à des ressources en littératie dans la langue d'éducation serait notamment facilité pour les élèves d'écoles anglophones du Québec à l'échelle canadienne et nord-américaine en comparaison aux élèves d'écoles francophones minoritaires ailleurs au Canada. Cette situation peut contribuer à créer des disparités, notamment dans les matières scolaires liées à la langue, entre les élèves des écoles de langue minoritaire au niveau des déterminants du processus éducatif. Malgré cet avantage par rapport aux élèves d'écoles francophones minoritaires du Canada hors Québec, les élèves d'écoles anglophones du Québec ont quand même des résultats inférieurs à ceux des élèves d'écoles francophones majoritaires du Québec en raisonnement mathématique et ce, en dépit de caractéristiques plus favorables.

Enfin, les résultats montrent que la proportion relative de l'écart inexpliqué par rapport à l'écart total est plus importante dans les provinces où les francophones minoritaires sont moins nombreux et moins concentrés géographiquement. Lors du recensement de 2006, la proportion de personnes qui ont déclaré le français comme première langue officielle parlée était de 4.5 % pour l'Ontario, 1.9 % pour l'Alberta et 1.5 % en Colombie-Britannique (Corbeil et Lafrenière, 2010 ;

Chavez *et al.*, 2011 ; Chavez et Bouchard-Coulombe, 2011). L'Ontario est caractérisée par une forte concentration francophone dans les régions du Nord-Ouest, d'Ottawa et du Sud-Est où réside 62.6 % de la population francophone de la province. Dans les régions du Nord-Ouest et du Sud-Est respectivement, 23.4 % et 60.2 % des francophones habitent dans des municipalités où ils représentent plus de 50 % de la population totale. Du côté de l'Alberta, seulement 1.3 % des francophones de la province habitent dans une municipalité où ils représentent plus de 50 % de la population totale. De plus, 70 % des francophones de la province résident dans les régions métropolitaines de Calgary et d'Edmonton où ils représentent respectivement 1.7 % et 2.1 % de la population totale. En Colombie-Britannique, 60 % des francophones de la province résident dans les régions métropolitaines de Vancouver et Victoria où ils représentent respectivement 1.5 % et 1.9 % de la population totale. Aucune personne recensée en Colombie-Britannique n'habite dans une municipalité où les francophones représentent plus de 50 % de la population totale.

Dans l'ensemble, les résultats obtenus dans le cadre de ce mémoire appuient partiellement la présence d'un effet minoritaire. C'est donc dire que les élèves qui fréquentent des écoles de langue minoritaire font face à des difficultés particulières relativement au processus éducationnel et que ces difficultés contribuent à exacerber l'écart entre les élèves des écoles de langue minoritaire et de langue majoritaire.

La présente étude comporte cependant des limites quant à l'interprétation des données dans le cadre de l'Article 23 de la Charte des droits et libertés du Canada. En effet, les données publiques de l'étude PISA 2009 ne permettent pas de faire une distinction entre la langue dominante chez l'élève et la langue d'enseignement de l'école. Les groupes ont été créés sur la base de la langue d'enseignement de l'école que les élèves fréquentent. Or, il n'est pas exclu, au Canada hors Québec, que des parents francophones ou des familles exogames (parents de langues différentes) décident d'envoyer leur enfant à l'école anglophone et vice versa. Les données publiques n'incluent que la langue parlée à la maison et la langue du test qui correspond à la langue d'enseignement de l'école. La langue parlée à la maison ne dit cependant rien sur la composition linguistique de la famille. Une famille peut en effet parler anglais à la maison malgré le fait qu'un des deux parents soit francophone et que leurs enfants soient conséquemment des

enfants d'ayants droit en vertu de l'article 23 de la Charte.³¹ La prévalence des familles exogames, notamment dans les provinces de l'ouest du Canada, fait en sorte que la langue parlée à la maison ne permet pas d'identifier adéquatement l'ensemble des élèves constituant les enfants d'ayants droit.³² La langue du test représente à cet égard une variable qui permet d'identifier plus adéquatement les élèves de langue minoritaire. Afin d'identifier les élèves constituant des enfants d'ayants droit en vertu de l'article 23, seuls les élèves fréquentant une école publiquement subventionnée ont été considérés dans l'analyse. De plus, le problème d'identification est grandement mitigé dans certaines provinces où l'admissibilité à l'école dans la langue de la minorité est restreinte aux enfants d'ayants droit définis par l'article 23.³³ D'un autre côté, il est beaucoup plus difficile d'identifier les enfants d'ayants droit qui fréquentent une école de langue majoritaire. Le Tableau 8 montre cependant que la proportion d'élèves inscrits dans une école de langue majoritaire qui parle la langue de la minorité linguistique à la maison est très faible, soit 0.4 % pour le Canada hors Québec et 2.77 % pour le Québec. De plus, les résultats des estimations des fonctions de production d'éducation présentés dans la section précédente montrent que le fait de ne pas parler la langue du test à la maison est associé significativement à des résultats inférieurs. On devrait donc s'attendre à ce que l'inclusion d'élèves qui parlent la langue de la minorité à la maison dans le groupe des élèves de langue majoritaire contribue à réduire l'écart entre les deux groupes. À ce titre, cette inclusion des élèves qui parlent la langue de la minorité à la maison dans le groupe des élèves de langue majoritaire constitue une mesure conservatrice.

À la lumière de ces considérations, la langue du test, et par conséquent celle de l'école, a été utilisée afin d'identifier les élèves de langue minoritaire et les élèves de langue majoritaire. Cependant, puisque les groupes ne sont pas homogènes en ce qui a trait à la langue ou aux langues parlées à la maison, et que du même fait les élèves de langue minoritaire ne sont pas tous

³¹ Les enfants d'ayants droit en vertu de l'article 23 sont les élèves dont un des deux parents a pour langue maternelle la langue de la minorité linguistique de la province dans laquelle il réside. De plus, l'article 23 s'applique aussi aux élèves dont l'un des deux parents, ou l'un des frères ou sœurs a reçu une éducation primaire dans la langue de la minorité.

³² Une famille exogame est définie par le fait que les deux conjoints n'ont pas la même langue maternelle. En Alberta (Colombie-Britannique), lors du recensement de 2006, 23 % (26 %) des enfants nés dans une famille exogame anglais-français se sont vu transmettre le français comme langue maternelle (Chavez *et al.*, 2011 ; Chavez et Bouchard-Coulombe, 2011).

³³ Plusieurs provinces canadiennes, notamment la Nouvelle-Écosse, le Québec, l'Ontario, le Manitoba, l'Alberta et la Colombie-Britannique restreignent l'accès à l'éducation publique dans la langue de la minorité. Des variantes provinciales existent quant à la définition de ce qui constitue un ayant droit, cependant, l'ensemble des définitions provinciales inclut la définition donnée par la Charte.

des enfants d'ayants droit en vertu de l'article 23, il est préférable de parler d'élèves d'écoles de langue minoritaire plutôt que d'élèves issus de la minorité linguistique.

Une limite additionnelle tient à la nature même des données et des variables utilisées dans les modèles. Les problèmes d'endogénéité relatifs à l'utilisation de variables approximatives font en sorte que l'interprétation des coefficients estimés doit être faite avec précaution. Malgré l'utilisation de variables qui tentent de contrôler pour des caractéristiques non observées, notamment l'indice d'attitude envers l'école et le nombre de livres à la maison, il demeure que l'éducation est un phénomène qui doit être compris dans sa multidimensionnalité. Ainsi, même si le fait de posséder plus de 100 livres à la maison est associé à des résultats moyens plus élevés, il ne suffit pas d'augmenter le nombre de livres dans la maison des élèves d'une école afin d'augmenter les résultats moyens dans l'école. La valeur des résultats repose davantage sur leur pouvoir descriptif que sur leur pouvoir normatif. Hanushek (1986) note aussi une limite des fonctions de production d'éducation avec données *cross*-sectionnelles en soulevant le fait que le processus éducationnel est un processus continu. Or, l'utilisation de données *cross*-sectionnelles ne prend pas en considération la variation des intrants dans le temps et du même fait, comment ces intrants passés influencent les résultats présents. À ce titre, l'utilisation de données panel permettrait de cibler plus adéquatement les déterminants du processus éducatif.

8. Conclusion

Le présent mémoire avait pour but de répondre à deux questions. Premièrement, *est-ce que les élèves d'écoles de langue minoritaire du Canada jouissent d'une égalité réelle en matière de prestation de services dans la langue de la minorité conformément à l'article 23 de la Charte?* Deuxièmement, *dans la mesure où il existe un écart en termes de résultats entre les élèves qui fréquentent une école de langue minoritaire et ceux qui fréquentent une école de langue majoritaire, qu'est-ce qui conditionne cet écart?* Sur la base de la langue de l'école fréquentée, les résultats montrent qu'il existe des écarts statistiquement significatifs entre les élèves d'écoles de langue minoritaire et majoritaire en ce qui a trait aux résultats sur les tests d'aptitudes PISA de compréhension en lecture et de raisonnement mathématique. À ce titre les élèves de langue minoritaire ne jouissent donc pas d'une égalité réelle en matière de prestation de services dans la langue de la minorité.

Les résultats de la décomposition Blinder-Oaxaca montrent aussi que, bien qu'une proportion de l'écart observé soit expliquée par des différences au niveau des caractéristiques individuelles, familiales, et institutionnelles, la majorité de l'écart observé n'est pas expliquée par des différences de caractéristiques incluses dans la fonction de production d'éducation. Cet écart inexpliqué est notamment la résultante de retours différents sur les caractéristiques en termes de résultats entre les élèves des écoles de langue minoritaire et majoritaire ainsi que de différences au niveau des caractéristiques non observées. Nous utilisons le terme *effet minoritaire* afin de définir les difficultés particulières aux élèves des écoles de langue minoritaire qui composent l'écart inexpliqué.

La pertinence de l'exercice effectué dans le présent mémoire repose sur deux prémisses. La première prémisses, d'ordre juridique, est la question de l'égalité réelle garantie en vertu de l'article 23 de la Charte. Dans l'arrêt *Mahé c Alberta*, la Cour Suprême du Canada a déclaré que « [...] la qualité de l'enseignement dispensé à la minorité linguistique devrait en principe être, dans une mesure raisonnable, égale à celle de l'enseignement donné à la majorité, sans avoir à être identique, et des fonds publics adéquats à cette fin doivent être fournis » (*Mahé c Alberta*, 1990). Or, nos résultats montrent que cette disposition n'est pas respectée, en moyenne, à l'échelle du pays.

La deuxième prémisses, d'ordre économique, repose sur les résultats des travaux de Behrman et Birdsall (1983) et de Neal et Johnson (1994). Ces auteurs ont trouvé que la qualité de l'éducation reçue avait un effet statistiquement significatif sur les revenus et qu'un écart de qualité pouvait expliquer une très grande proportion de l'écart salarial entre deux groupes. Ainsi, il n'est pas exclu que les écarts observés dans le cadre du présent mémoire puissent se traduire à long terme par des écarts salariaux entre les minorités et majorités linguistiques.³⁴

Les résultats présentés dans le cadre de ce mémoire ne permettent pas de déterminer l'action appropriée afin d'améliorer les résultats des élèves de langue minoritaire, mais ils montrent néanmoins qu'il est impératif pour les décideurs publics d'agir sur la question. La présence d'un effet minoritaire doit être prise en compte dans le calcul des ressources allouées aux écoles de langue minoritaire. À ce titre, il pourrait être intéressant de procéder à des analyses de différence-en-différence pour voir dans quelle mesure l'octroi de ressources additionnelles et

³⁴ Lepage (2012) montre que les francophones minoritaires ont des salaires en moyenne plus élevés que les anglophones majoritaires en Ontario et dans les provinces de l'Ouest. Cependant, l'étude de Christofides et Swidinsky (2010) montre que les salaires plus élevés des francophones minoritaires pourraient être dus à leur meilleure connaissance de la deuxième langue officielle ainsi qu'à une plus grande demande pour de la main-d'œuvre bilingue dans la fonction publique.

la façon dont elles sont utilisées dans le cadre du processus éducatif ont un effet sur les résultats des élèves et dans quelle mesure cet octroi permet de réduire les écarts observés entre les élèves d'écoles de langue minoritaire et majoritaire.

Références

Références académiques

Ammermueller, A. (2007a). Poor background or low returns? Why immigrant students in Germany perform so poorly in the programme for international student assessment. *Education Economics*, 15(2), 215-230.

Ammermueller, A. (2007b). PISA: What makes the difference? *Empirical Economics*, 33(2), 263-287.

Behrman, J. R., & Birdsall, N. (1983). The quality of schooling: quantity alone is misleading. *The American Economic Review*, 73(5), 928-946.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human Resources*, 8(4), 436-455.

Breton, A. (1999). Le rendement culturel des langues et l'assimilation linguistique, dans A. Breton (dir.). *Explorer l'économie linguistique*, Travaux publics et Services gouvernementaux Canada, [Nouvelles perspectives canadiennes], Ottawa, 95-121.

Chavez, B. & Bouchard-Coulombe, C. (2011). Portrait des minorités de langue officielle au Canada : les francophones de la Colombie-Britannique. *Statistique Canada, Division de la Statistique Sociale et Autochtone*, (numéro au catalogue 89-642-X-004), disponible à l'adresse <http://www.statcan.gc.ca/pub/89-642-x/89-642-x2011004-fra.pdf>

Chavez, B., Bouchard-Coulombe, C. & Lepage, J.-F. (2011). Portrait des minorités de langue officielle au Canada : les francophones de l'Alberta. *Statistique Canada, Division de la Statistique Sociale et Autochtone*, (numéro au catalogue 89-642-X-007), disponible à l'adresse <http://www.statcan.gc.ca/pub/89-642-x/89-642-x2011007-fra.pdf>

Coleman, J. S., Campbell E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966) Equality of educational opportunity. *National Center for Educational Statistics*, Washington DC.

Commissariat aux langues officielles. (2011). Les langues officielles de la population du Canada. Dernière date de modification: 15-08-2011, disponible à l'adresse http://www.ocolclo.gc.ca/html/provinces_f.php

Corbeil, J.-P. & Lafrenière, S. (2010). Portrait des minorités de langue officielle au Canada : les francophones de l'Ontario. *Statistique Canada, Division de la Statistique Sociale et Autochtone*, (numéro au catalogue 89-642-X-001), disponible à l'adresse <http://www.statcan.gc.ca/pub/89-642-x/89-642-x2010001-fra.pdf>

Chingos, M. M., & Peterson, P. E. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review*, 30(3), 449-465.

- Christofides, L. N., & Swidinsky, R. (2010). The economic returns to the knowledge and use of a second official language: English in Quebec and French in the rest-of-Canada. *Canadian Public Policy*, 36(2), 137-158.
- Corak, M., & Lauzon, D. (2009). Differences in the distribution of high school achievement: The role of class-size and time-in-term. *Economics of Education Review*, 28(2), 189-198.
- Currie, J., & Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4), 1495-1532.
- Duque, J. C., Ramos R., & Nieto S. (2012). Decomposing the rural-urban differential in student achievement in Colombia using PISA microdata. *IZA Discussion Paper*, No. 6515, Avril 2012.
- Fryer Jr, R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics*, 86(2), 447-464.
- Fuchs, T., & Woessmann, L. (2007). What accounts for international differences in student performance? A re-examination using PISA data. *Empirical Economics*, 32(2), 433-464.
- Grenier, G., & Lacroix, G. (1986). Les revenus et la langue: le cas de la capitale nationale. *L'Actualité Économique*, 62(3), 365-384.
- Hanushek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24(3), 1141-1177.
- Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review*, 90(5), 1184-1208.
- Hanushek, E. A., & Luque, J. A. (2003). Efficiency and equity in schools around the world. *Economics of Education Review*, 22(5), 481-502.
- Hanushek, E. A. (2005). Why quality matters in education. *Finance and Development*, 42(2), 15-19.
- Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3), 607-668.
- Hanushek, E. & Woessmann, L. (2010). The economics of international differences in educational achievement. *IZA Discussion Paper*, No. 4925.
- Hanushek, E.A. (2011). The economic value of higher teacher quality. *Economics of Education Review*, 30(3), 466-479.
- Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4), 453-479.

- Judkins, D. R. (1990). Fay's method for variance estimation. *Journal of Official Statistics*, 6(3), 223-239.
- Juhn, C., Murphy, K. M., & Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3), 410-442.
- Lavoie, M., & Saint-Germain, M. (1991). Disparités linguistiques de revenu au Canada selon la langue parlée à la maison. *L'Actualité Économique*, 67(3), 356-380.
- Lee, J. W., & Barro, R. J. (2001). Schooling quality in a cross-section of countries. *Economica*, 68(272), 465-488.
- Leigh, A., & Gong, X. (2009). Estimating cognitive gaps between Indigenous and non-Indigenous Australians. *Education Economics*, 17(2), 239-261.
- Lepage, J.-F. (2012). Situation des minorités de langue officielle sur le marché du travail. *Statistique Canada, Division de la Statistique Sociale et Autochtone*, (numéro au catalogue 89-651-X2012001), disponible à l'adresse <http://www.statcan.gc.ca/pub/89-651-x/89-651-x2012001-fra.pdf>
- Macdonald, K. (2008). PV: Stata module to perform estimation with plausible values. *Statistical Software Components from Boston College Department of Economics*, dernière date de modification: 12-04-2011, disponible à l'adresse <http://econpapers.repec.org/software/bocbo/code/s456951.htm>
- McEwan, P. J., & Marshall, J. H. (2004). Why does academic achievement vary across countries? evidence from Cuba and Mexico. *Education Economics*, 12(3), 205-217.
- Meunier, M. (2011). Immigration and student achievement: Evidence from Switzerland. *Economics of Education Review*, 30(1), 16-38.
- Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56(2), 177-196.
- Neal, D. A., & Johnson, W. R. (1996). The role of premarket factors in black-white wage differences. *Journal of Political Economy*, 104(5), 869.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3), 693-709.
- OCDE. (2012). PISA 2009 Technical Report. OCDE, Paris, disponible à l'adresse <http://www.oecd.org/pisa/pisaproducts/pisa2009/50036771.pdf>
- OCDE. (2009). PISA Data Analysis Manual: SAS Second Edition. OCDE, Paris, disponible à l'adresse <http://browse.oecdbookshop.org/oecd/pdfs/free/9809021e.pdf>

Sakellariou, C. (2008). Peer effects and the indigenous/non-indigenous early test-score gap in Peru. *Education Economics*, 16(4), 371-390.

Schneeweis, N. (2011). Educational institutions and the integration of migrants. *Journal of Population Economics*, 24(4), 1281-1308.

Schnepf, S. V. (2007). Immigrants' educational disadvantage: an examination across ten countries and three surveys. *Journal of Population Economics*, 20(3), 527-545.

Shapiro, D. M., & Stelcner, M. (1997). Language and earnings in Quebec: trends over twenty years, 1970-1990. *Canadian Public Policy*, 23(2) 115-140.

Shipley, L. (2011). A profile of minority-language students and schools in Canada: Results from the Programme for International Student Assessment (PISA), 2009. *Culture, Tourism and the Centre for Education Statistics Research Papers*, disponible à l'adresse: <http://www.statcan.gc.ca/pub/81-595-m/81-595-m2011092-eng.pdf>

Willms, D. (2004). Variation in literacy skills among Canadian provinces: Findings from the OECD PISA. *Culture, Tourism and the Centre for Education Statistics Research papers*, disponible à l'adresse: <http://publications.gc.ca/Collection/Statcan/81-595-MIE/81-595-MIE2004012.pdf>

Woessmann, L. (2003). Schooling resources, educational institutions and student performance: the international evidence. *Oxford Bulletin of Economics and Statistics*, 65(2), 117-170.

Woessmann, L., & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, 50(3), 695-736.

Woessmann, L. (2011). Cross-country evidence on teacher performance pay. *Economics of Education Review*, 30(3), 404-418.

Worswick, C. (2004). Adaptation and inequality: children of immigrants in Canadian schools. *Canadian Journal of Economics*, 37(1), 53-77.

Références juridiques

Arsenault-Cameron c Île-du-Prince-Édouard, [2000] 1 RCS 3.

Charte canadienne des droits et libertés, partie I de la *Loi constitutionnelle de 1982*, constituant l'annexe B de la *Loi de 1982 sur le Canada* (R-U), 1982, c 11.

Mahé c Alberta, [1990] 1 RCS 342.

Annexe 1 : Description des variables d'intérêts

Niveau de l'élève

Résultats de tests de compréhension en lecture et de raisonnement mathématique. Les résultats sont définis par cinq valeurs plausibles pour chacun des tests et pour chaque étudiant. Les valeurs plausibles sont standardisées de façon à avoir une moyenne internationale de 500 et un écart-type international de 100.

Sexe. Cette variable a été dichotomisée de façon à avoir les garçons comme catégorie de référence.

Cheminement normal. Cette variable rend compte du parcours scolaire de l'élève. La variable a été dichotomisée de façon à avoir d'une part les élèves qui accusent un retard sur le cheminement normal (catégorie de référence) et d'autre part les élèves qui n'accusent aucun retard, ou qui sont en avance sur le cheminement normal. Essentiellement cette variable sert à identifier les élèves qui ont échoué une année scolaire ou plus dans le passé. Cependant, les réglementations particulières du Québec sur l'âge requis pour débiter l'école font en sorte que les élèves qui sont nés après le 1^{er} octobre doivent commencer l'école une année plus tard. Ces élèves sont comptabilisés comme accusant une année de retard relativement au cheminement normal. Même si les raisons qui motivent cette année de retard chez les élèves nés après le 1^{er} octobre ne sont pas les mêmes que les élèves qui ont échoué, il demeure que les élèves nés après le 1^{er} octobre sont en retard d'une année sur le programme de la majorité des élèves de 15 ans. De plus, puisque les élèves anglophones et francophones sont soumis aux mêmes conditions au Québec et que la proportion des élèves nés dans les trois derniers mois de l'année est relativement la même dans les deux groupes, la différence entre les deux groupes au niveau de la variable découle des élèves qui ont échoué une année dans le passé. Cette variable ne peut cependant pas être utilisée aux fins de comparaisons entre le Québec et les autres provinces du Canada.

Immigrant. Cette variable a été dichotomisée de façon à avoir d'une part les élèves qui sont nés au Canada (catégorie de référence) et d'autre part les élèves qui sont nés à l'extérieur du Canada.

Indice d'attitude envers l'école. Cette variable est construite par le consortium PISA suivant les spécifications de la théorie des réponses aux items et plus particulièrement les spécifications du modèle de Rasch exposé dans la section précédente. Quatre questions répondues par l'élève sont utilisées pour construire cet indice. Le score propre à chaque élève est obtenu avec l'estimation

du maximum de vraisemblance pondéré qui reflète la probabilité de répondre par la négative aux questions. Les scores sont ensuite normalisés pour avoir une moyenne internationale de 0 et un écart type international de 1. Une valeur positive de l'indice rend compte d'une attitude envers l'école plus positive que la moyenne internationale et une valeur négative de l'indice rend compte d'une attitude envers l'école plus négative que la moyenne internationale.

Niveau familial

Langue parlée à la maison. Cette variable rend compte de la langue la plus souvent parlée à la maison. Elle a été dichotomisée de façon à avoir d'une part ceux dont la langue parlée le plus souvent à la maison est une autre langue que la langue du test (catégorie de référence) et d'autre part ceux dont la langue parlée le plus souvent à la maison est la langue du test.

Structure familiale. Cette variable a été dichotomisée afin d'avoir d'une part les élèves issus d'une famille non traditionnelle (ceux qui n'ont pas deux parents, biologiques ou non, qui habitent sous le même toit) et d'autre part les élèves issus d'une famille traditionnelle (ceux qui ont deux parents, biologiques ou non, qui habitent sous le même toit).

*Éducation de la mère.*³⁵ Cette variable a été dichotomisée de façon à avoir d'une part les élèves dont la mère ne possède pas un diplôme universitaire (catégorie de référence), c'est-à-dire un baccalauréat, une maîtrise, ou un doctorat, et d'autre part les élèves dont la mère possède un diplôme universitaire.

Nombre de livres à la maison. Cette variable a été dichotomisée de façon à avoir d'une part les élèves qui ont rapporté avoir 100 livres et moins à la maison (catégorie de référence) et d'autre part les élèves qui ont rapporté avoir plus de 100 livres à la maison.

Niveau institutionnel

Population dans la communauté. Cette variable a été dichotomisée de façon à avoir d'une part les élèves qui fréquentent une école qui se situe dans une communauté de moins de 15 000 habitants (catégorie de référence) et d'autre part les élèves qui fréquentent une école qui se situe dans une communauté de 15 000 habitants et plus.

³⁵ L'inclusion de la variable *éducation de la mère* par opposition à *éducation du père* est notamment justifiée par les travaux de Currie et Moretti (2003)

Ratio d'ordinateurs par élève dans l'école. Cette variable a été dichotomisée de façon à avoir d'une part les élèves qui fréquentent une école où il y a moins d'un ordinateur par élève (catégorie de référence) et d'autre part les élèves qui fréquentent une école où il y a un ordinateur ou plus par élève.

Ratio d'élèves par enseignant. Cette variable a été dichotomisée de façon à avoir d'une part les élèves qui fréquentent une école où il y a 17 élèves ou moins pour chaque enseignant (catégorie de référence) et d'autre part les élèves qui fréquentent une école où il y a plus de 17 élèves pour chaque enseignant.

Nombre d'étudiants dans l'école. Cette variable a été dichotomisée de façon à avoir d'une part les élèves qui fréquentent une école de moins de 1 000 élèves (catégorie de référence) et d'autre part les élèves qui fréquentent une école de 1 000 élèves et plus.

Indice du niveau socio-économique moyen dans l'école. Cette variable est utilisée pour capturer l'effet des pairs.³⁶ L'indice du niveau socio-économique, construit par le consortium PISA, est un indice agrégé de trois indices, soit l'indice du plus haut niveau d'éducation atteint par les parents, l'indice du plus haut niveau d'emploi atteint par les parents et l'indice de richesse culturelle à la maison. Les scores sur l'indice socio-économique sont normalisés pour avoir une moyenne internationale de 0 et un écart type international de 1. La variable est la moyenne de l'indice socio-économique dans l'école fréquentée par l'élève. Une valeur positive de la variable rend compte d'un niveau socio-économique moyen dans l'école qui est supérieur au niveau socio-économique moyen international et une valeur négative de la variable rend compte d'un niveau socio-économique moyen dans l'école qui est inférieur au niveau socio-économique moyen international.

Indice de l'attitude des enseignants. Cette variable est construite par le consortium PISA suivant les spécifications de la théorie des réponses aux items et plus particulièrement les spécifications du modèle de Rasch exposé dans la section précédente. Sept questions répondues par le directeur de l'école sont utilisées pour construire cet indice. Le score propre à chaque école est obtenu avec l'estimation du maximum de vraisemblance pondéré qui reflète la probabilité de répondre par la négative aux questions. Les scores sont ensuite normalisés pour avoir une moyenne internationale

³⁶ Duque *et al.* (2012) utilisent le niveau socio-économique moyen dans l'école pour rendre compte de l'effet de pairs. D'autres variables, évaluées à leur valeur moyenne au niveau de l'école, peuvent aussi être utilisées afin de rendre compte d'un effet de pairs, notamment les résultats moyens de l'école sur les tests d'aptitudes (Sakellariou, 2008).

de 0 et un écart type international de 1. Une valeur positive de l'indice rend compte d'une attitude des enseignants envers leur travail qui est plus positive que la moyenne internationale et une valeur négative de l'indice rend compte d'une attitude des enseignants envers leur travail qui est plus négative que la moyenne internationale.

Tableaux

Tableau 1 : Fréquences et pourcentages de la répartition des écoles publiques et du nombre d'élèves par région et par groupe linguistique dans l'échantillon final

	Canada hors Québec			Québec		
	Minorité linguistique (francophone)	Majorité linguistique (anglophone)	Total	Minorité linguistique (anglophone)	Majorité linguistique (francophone)	Total
Écoles (N = 681)	120 (23.17 %)	398 (76.83 %)	518	53 (32.5 %)	111 (67.5 %)	163
Élèves (N = 15 360)	3 153 (25.74 %)	9 095 (74.26 %)	12 248	945 (30.4 %)	2 167 (69.6 %)	3 112

Minorité (majorité) linguistique désigne les écoles de langue minoritaire (majoritaire) et les élèves d'école de langue minoritaire (majoritaire).

Tableau 2 : Moyennes pondérées et erreurs standards des variables par région géographique et groupe linguistique

	Canada hors Québec		Québec	
	Minorité linguistique (francophone)	Majorité linguistique (anglophone)	Minorité linguistique (anglophone)	Majorité linguistique (francophone)
	N = 2 234	N = 6 983	N = 703	N = 1 731
Résultat compréhension lecture	483.59 (2.17)	533.76 (2.33)	522.44 (3.72)	529.79 (3.89)
Résultat raisonnement mathématique	504.96 (2.21)	528.27 (2.42)	537.55 (4.30)	551.05 (4.13)
Sexe	0.53 (0.003)	0.51 (0.003)	0.51 (0.01)	0.52 (0.01)
Cheminement normal	0.91 (0.003)	0.94 (0.003)	0.68 (0.01)	0.60 (0.01)
Immigrant	0.06 (0.003)	0.14 (0.01)	0.06 (0.01)	0.06 (0.01)
Indice d'attitude envers l'école	0.23 (0.01)	0.06 (0.01)	-0.07 (0.03)	0.22 (0.02)
Langue parlée à la maison	0.57 (0.01)	.86 (0.01)	0.76 (0.01)	0.90 (0.01)
Structure familiale	0.84 (0.004)	0.82 (0.003)	0.82 (0.01)	0.82 (0.004)
Éducation de la mère	0.37 (0.01)	0.35 (0.004)	0.31 (0.01)	0.30 (0.01)
Nombre de livres à la maison	0.39 (0.01)	0.51 (0.01)	0.54 (0.01)	0.35 (0.01)
Population dans la communauté	0.59 (0.01)	0.79 (0.01)	0.87 (0.01)	0.74 (0.02)
Ratio d'ordinateurs/élève dans l'école	0.52 (0.01)	0.28 (0.02)	0.11 (0.01)	0.13 (0.02)
Ratio d'élèves/enseignant	0.04 (0.002)	0.29 (0.01)	0.60 (0.01)	0.43 (0.02)
Nombre d'élèves dans l'école	0.14 (0.002)	0.54 (0.01)	0.48 (0.01)	0.66 (0.02)
Indice de l'attitude des enseignants	-0.15 (0.004)	-0.04 (0.02)	-0.19 (0.01)	-0.34 (0.04)
Indice du niveau socio-économique moyen dans l'école	0.47 (0.002)	0.53 (0.01)	0.52 (0.01)	0.35 (0.02)

Minorité (majorité) linguistique désigne les élèves d'école de langue minoritaire (majoritaire). Les valeurs en caractère gras désignent un écart significatif (5 %) entre les deux groupes de la même région. Les strates ne contenant qu'une unité d'échantillonnage sont identifiées comme des unités d'échantillonnages non probabilistes et ne contribuent donc pas à l'échantillonnage de la variance.

Tableau 3 : Estimation de la fonction de production d'éducation en compréhension en lecture pour les élèves de langue minoritaire et majoritaire du Canada hors Québec et du Québec

	Canada hors Québec		Québec	
	Minorité linguistique (francophone)	Majorité linguistique (anglophone)	Minorité linguistique (anglophone)	Majorité linguistique (francophone)
	N = 2 234	N = 6 983	N = 703	N = 1 731
Sexe	27.73*** (4.10)	27.63*** (2.63)	21.00*** (6.77)	20.18*** (3.70)
Cheminement normal	59.94*** (6.42)	50.56*** (6.44)	35.45*** (8.96)	47.73*** (4.33)
Immigrant	-8.99 (10.43)	0.71 (4.58)	-20.19 (15.00)	-2.83 (11.07)
Indice d'attitude envers l'école	4.20** (1.77)	9.21*** (1.36)	5.61* (3.21)	6.80*** (1.83)
Langue parlée à la maison	13.18*** (4.19)	6.60 (5.87)	24.42*** (8.24)	18.02*** (6.81)
Structure familiale	11.79** (5.69)	8.20** (3.49)	6.23 (8.52)	-5.07 (5.54)
Éducation de la mère	11.50*** (4.06)	13.92*** (2.76)	-4.71 (7.94)	0.44 (4.31)
Nombre de livres à la maison	33.70*** (4.29)	38.69*** (2.32)	37.19*** (7.49)	32.27*** (4.61)
Population dans la communauté	9.21** (4.42)	10.05** (4.51)	20.81* (11.12)	7.22 (5.97)
Ratio d'ordinateurs/élève dans l'école	4.02 (3.53)	3.73 (4.20)	3.99 (12.09)	-5.51 (6.32)
Ratio d'élèves/enseignant	27.08* (14.92)	0.24 (4.09)	22.37*** (6.79)	7.03 (5.52)
Nombre d'élèves dans l'école	15.63** (6.45)	-0.32 (4.31)	-19.06** (8.81)	0.31 (5.65)
Indice de l'attitude des enseignants	5.44* (2.90)	1.37 (3.09)	4.50 (3.44)	-1.36 (3.96)
Indice du niveau socio- économique moyen dans l'école	25.66*** (9.23)	38.80*** (6.21)	46.63*** (11.14)	45.47*** (9.33)
Constante	356.97*** (8.42)	405.68*** (9.38)	400.81*** (13.85)	441.85*** (9.13)
R ² moyen	.1946	.1932	.2396	.2814

Minorité (majorité) linguistique désigne les élèves d'école de langue minoritaire (majoritaire). *, ** et *** signifient que les résultats sont significatifs à 10%, 5% et 1% respectivement.

Tableau 4 : Estimation de la fonction de production d'éducation en raisonnement mathématique pour les élèves de langue minoritaire et majoritaire du Canada hors Québec et du Québec

	Canada hors Québec		Québec	
	Minorité linguistique (francophone)	Majorité linguistique (anglophone)	Minorité linguistique (anglophone)	Majorité linguistique (francophone)
	N = 2 234	N = 6 983	N = 703	N = 1 731
Constante	411.35*** (7.78)	436.02*** (10.05)	436.85*** (16.05)	478.79*** (11.36)
Sexe	-19.24*** (3.64)	-17.12*** (2.54)	-20.53*** (7.29)	-27.82*** (3.47)
Cheminement normal	65.69*** (5.74)	49.77*** (7.92)	46.90*** (8.69)	61.73*** (4.59)
Immigrant	-17.74* (10.53)	-2.04 (4.66)	-13.02 (13.22)	-1.33 (12.19)
Indice d'attitude envers l'école	2.53 (2.18)	6.74*** (1.46)	4.86 (3.70)	5.49*** (1.84)
Langue parlée à la maison	-3.23 (4.25)	-8.59 (5.35)	11.09 (9.34)	16.60* (8.96)
Structure familiale	15.77** (5.86)	9.79*** (3.38)	16.58* (9.25)	4.04 (5.25)
Éducation de la mère	14.38*** (3.96)	13.99*** (2.62)	5.33 (9.40)	6.70 (4.47)
Nombre de livres à la maison	35.42*** (4.60)	38.92*** (3.10)	40.00*** (6.82)	30.65*** (4.97)
Population dans la communauté	-1.99 (4.98)	5.53 (5.51)	27.28** (12.84)	-2.49 (7.92)
Ratio d'ordinateurs/élève dans l'école	1.14 (3.91)	7.34 (4.50)	-5.44 (14.46)	-2.30 (7.19)
Ratio d'élèves/enseignant	12.14 (14.94)	4.35 (4.24)	4.94 (7.16)	1.05 (6.59)
Nombre d'élèves dans l'école	9.96 (6.51)	1.80 (4.76)	-17.56** (8.81)	-1.39 (6.92)
Indice de l'attitude des enseignants	1.53 (3.06)	3.65 (2.50)	7.87** (3.60)	0.68 (4.62)
Indice du niveau socio- économique moyen dans l'école	27.68** (11.26)	38.88*** (6.69)	36.93*** (13.80)	58.59*** (11.28)
R ² moyen	.1946	.1764	.2354	.34

Minorité (majorité) linguistique désigne les élèves d'école de langue minoritaire (majoritaire). *, ** et *** signifient que les résultats sont significatifs à 10%, 5% et 1% respectivement.

Tableau 5 : Décomposition Blinder-Oaxaca de l'écart des résultats en compréhension en lecture entre les élèves de langue minoritaire et majoritaire du Canada hors Québec et du Québec

	Canada hors Québec (N = 9 217)	Québec (N = 2 434)
Résultat moyen compréhension lecture : Majorité linguistique	533.76 (2.32)	529.79 (3.89)
Résultat moyen compréhension lecture : Minorité linguistique	483.59 (2.17)	522.44 (3.72)
Écart total (Majorité – Minorité)	50.16 (3.32)***	7.35 (5.52)
Écart inexpliqué	41.68 (3.46)***	22.32 (4.94)***
Écart expliqué	8.48 (2.77)***	-14.97 (3.71)***
Décomposition de l'écart expliqué		
Sexe	-0.75 (0.28)***	0.11 (0.36)
Cheminement normal	1.49 (0.53)***	-3.71 (1.18)***
Immigrant	0.06 (0.36)	-0.02 (0.11)
Indice d'attitude envers l'école	-1.57 (0.36)***	1.99 (0.68)***
Langue parlée à la maison	1.87 (1.70)	2.50 (0.98)**
Structure familiale	-0.20 (0.11)*	0.03 (0.13)
Éducation de la mère	-0.29 (0.20)	0.004 (0.06)
Nombre de livres à la maison	4.39 (0.67)***	-6.26 (1.18)***
Population dans la communauté	2.01 (0.95)**	-0.95 (0.88)
Ratio d'ordinateurs/élève dans l'école	-0.91 (1.04)	-0.11 (0.26)
Ratio d'élèves/enseignant	0.06 (1.04)	-1.22 (1.08)
Nombre d'élèves dans l'école	-0.13 (1.72)	0.06 (1.02)
Indice de l'attitude des enseignants	0.15 (0.31)	0.20 (0.63)
Indice du niveau socio-économique moyen dans l'école	2.28 (0.94)**	-7.58 (2.05)***

Minorité (majorité) linguistique désigne les élèves d'école de langue minoritaire (majoritaire). *, ** et *** signifient que les résultats sont significatifs à 10%, 5% et 1% respectivement.

Tableau 6 : Décomposition Blinder-Oaxaca de l'écart des résultats en raisonnement mathématique entre les élèves de langue minoritaire et majoritaire du Canada hors Québec et du Québec

	Canada hors Québec (N = 9 217)	Québec (N = 2 434)
Résultat moyen compréhension lecture : Majorité linguistique	528.27 (2.42)	551.05 (4.13)
Résultat moyen compréhension lecture : Minorité linguistique	504.96 (2.21)	537.55 (4.30)
Écart total (Majorité – Minorité)	23.31 (3.30)***	13.50 (5.62)**
Écart inexpliqué	17.43 (3.67)***	30.60 (5.75)***
Écart expliqué	5.88 (2.95)**	-17.10 (4.46)***
Décomposition de l'écart expliqué		
Sexe	0.46 (0.16)***	-0.16 (0.48)
Cheminement normal	1.47 (0.58)**	-4.79 (1.54)***
Immigrant	-0.17 (0.38)	-0.01 (0.12)
Indice d'attitude envers l'école	-1.15 (0.32)***	1.60 (0.61)**
Langue parlée à la maison	-2.44 (1.53)	2.30 (1.32)*
Structure familiale	-0.24 (0.13)	-0.02 (0.08)
Éducation de la mère	-0.29 (0.20)	-0.059 (0.17)
Nombre de livres à la maison	4.42 (0.70)***	-5.95 (1.16)***
Population dans la communauté	1.11 (1.12)	0.33 (1.08)
Ratio d'ordinateurs/élève dans l'école	-1.79 (1.15)	-0.05 (0.19)
Ratio d'élèves/enseignant	1.10 (1.10)	-0.18 (1.20)
Nombre d'élèves dans l'école	0.71 (1.92)	-0.25 (1.26)
Indice de l'attitude des enseignants	0.40 (0.32)	-0.10 (0.77)
Indice du niveau socio-économique moyen dans l'école	2.28 (0.92)**	-9.77 (2.59)***

Minorité (majorité) linguistique désigne les élèves d'école de langue minoritaire (majoritaire). *, ** et *** signifient que les résultats sont significatifs à 10%, 5% et 1% respectivement.

Tableau 7 : Décomposition Blinder-Oaxaca de l'écart des résultats en compréhension lecture entre les élèves de langue minoritaire et majoritaire de l'Ontario, de l'Alberta et de la Colombie-Britannique

	Ontario (N = 3 146)	Alberta (N = 1 581)	Colombie-Britannique (N = 1 617)
Résultat moyen compréhension lecture : Majorité linguistique	536.95 (3.29)	531.93 (6.82)	533.32 (4.04)
Résultat moyen compréhension lecture : Minorité linguistique	483.99 (2.81)	482.65 (8.01)	495.47 (13.76)
Écart total (Majorité – Minorité)	52.95*** (4.45)	49.28*** (10.23)	37.85** (14.30)
Écart expliqué	8.16 (5.02)	5.85 (13.55)	0.13 (9.94)
Écart inexpliqué	44.80*** (5.55)	43.43** (17.19)	37.72** (16.61)

Minorité (majorité) linguistique désigne les élèves d'école de langue minoritaire (majoritaire). *, ** et *** signifient que les résultats sont significatifs à 10%, 5% et 1% respectivement.

Tableau 8 : Pourcentages de la répartition des élèves qui fréquentent une école francophone et des élèves qui fréquentent une école de langue anglophone au Canada hors Québec selon la langue parlée à la maison

	Langue parlée à la maison : Français	Langue parlée à la maison : Anglais	Langue parlée à la maison : Autres langues
Élèves qui fréquentent une école francophone	56.28 %	38.54 %	5.18 %
Élèves qui fréquentent une école anglophone	0.40 %	85.99 %	13.61 %
Pourcentage de la population totale du Canada hors Québec	2.51 %	84.20 %	13.29 %