Sorting, peers and achievement of Aboriginal students in British Columbia[1]

Jane Friesen
Simon Fraser University
Brian Krauth
Simon Fraser University

August 2009

Abstract

We use administrative data on students in grades 4 and 7 in British Columbia to examine the extent to which differences in school environment contribute to the achievement gap between Aboriginal and non-Aboriginal students as measured by standardized test scores. We find that segregation of Aboriginal and non-Aboriginal students is substantial, and that differences in the distribution of these two groups across schools account for roughly half the overall achievement gap on the Foundation Skills Assessment tests in grade 7. The substantial school-level segregation of Aboriginal and non-Aboriginal student across schools means that Aboriginal students on average have a higher proportion of peers who are themselves Aboriginal, as well as a higher proportion of peers in special education. We estimate the effect of peer composition on value-added exam outcomes, using longitudinal data on multiple cohorts of students together with school-by-grade fixed effects to account for endogenous selection into schools. We find that having a greater proportion of Aboriginal peers, if anything, improves the achievement of Aboriginal students.

JEL Codes: I21
Keywords: Aboriginal education, peer effects

---

[1] Edudata Canada provided the data used in this research, assembled from B.C. Ministry of Education records by Maria Trache. Funding for this project provided by Human Resources and Skills Development Canada through the Canadian Skills and Labour Market Research Network and by Simon Fraser University's Community Trust Endowment Fund is gratefully acknowledged. Michele Battisti provided excellent research assistance.
Corresponding author: Jane Friesen (friesen@sfu.ca).

1 Introduction

Aboriginal Canadians have an above-average incidence of almost every marker of social and economic deprivation, including poverty (Mendelson 2006), poor health outcomes, drug and alcohol addiction, and suicide (Health Canada 2009). Some analysts (e.g. Richards and Vining 2004) argue that the key to breaking the cycle of poverty among off-reserve Aboriginal Canadians lies in improving educational outcomes among Aboriginal children and youth. This view is supported by evidence from other populations that education is associated with better health behaviours and outcomes (Kenkel 1991), substantially lower rates of incarceration (Lochner and Moretti 2004), higher earnings (Card 1999), reduced teen childbearing, criminal propensity, child abuse and neglect, and improved educational attainment and health outcomes of children (Greenwood 1997), increased voter and civic participation (Dee 2003), and reduced reliance on public transfers (Wolfe and Haveman 2001).

Our goal in this paper is to contribute to establishing an evidence base that can inform the development of policies related to Aboriginal education in Canada. We use a newly available administrative data set provided by the British Columbia (B.C.) Ministry of Education to document the achievement gap between Aboriginal and non-Aboriginal students in B.C. as measured by standardized test scores in grades 4 and 7, and to investigate the relationship between this gap and student characteristics, particularly differences in rates of assessed disabilities. We next measure the extent to which Aboriginal students are segregated from non-Aboriginal students at school. B.C.'s school funding rules provide districts with roughly similar resource levels, so this source of variation in school quality is not as salient as in the U.S. context.[2] However, if peer effects are important, differential sorting of Aboriginal and non-Aboriginal students may lead to systematic differences in the quality of the learning environments of Aboriginal and non-Aboriginal students. We provide econometric estimates of the effects of peer group composition on Aboriginal students' achievement as measured by

[2] Even with the same funding levels, schools that serve Aboriginal populations may have greater difficulty attracting and retaining teaching and administrative staff if they are geographically isolated.

B.C.'s Foundation Skills Assessment tests. We focus in particular on the share of peers who are Aboriginal or who are classified as disabled.

Our results show that the grade 7 test score gap is large in both reading and numeracy. Most of the gap has developed by grade 4, but the gap continues to grow between grades 4 and 7. We find that differences in rates of identified disability do not explain much of the test score gap. We find a substantial degree of segregation between Aboriginal and non-Aboriginal students, suggesting that school environments can in principle play an important role in the achievement gap. We decompose the mean grade 7 test score gap into between-school (the extent to which Aboriginal students tend to attend schools in which both Aboriginal and non-Aboriginal students do poorly on the exams) and within-school (the extent to which Aboriginal students do worse on the exam than non-Aboriginal students in the same school) gaps. We find that about half of the gap takes the form of between-school variation. Some of this between-school variation might be explained by peer effects. The average Aboriginal student has a substantially higher proportion of Aboriginal peers and a somewhat higher proportion of peers with disabilities. However, we do not find that the characteristics of Aboriginal students' peers that result from this sorting contribute significantly to the relatively low test scores of Aboriginal students. If anything, Aboriginal students perform better when they attend school with a greater proportion of peers who are themselves Aboriginal, and experience limited if any disadvantage from attending school with a greater proportion of peers with disabilities.

2 Data and institutional background

2.1 On-reserve and off-reserve schooling in B.C.

Education in Canada falls under provincial jurisdiction, with the important exception of on-reserve education. On-reserve schools are generally operated by First Nations bands, with funding from the federal Department of Indian and Northern Affairs Canada (INAC). Students attending band-run schools do not appear in our data, so our results describe only those Aboriginal students who attend provincial schools. Comparable data for students attending on-reserve schools do not exist, so it is not possible to know how their academic achievement compares to Aboriginal students attending provincial schools. The evidence that exists suggests that on-reserve students have worse academic outcomes than off-reserve students do. Richards et al. (2008) use Census data to estimate that the high school completion rate of 20-24 year-olds living on reserve is 22 percentage points lower than that of First Nations 20-24 year-olds living off reserve (39% versus 62%), and 36 percentage points lower than that of Métis 20-24 year-olds (75%). The lower rates of high school completion among First Nations living on-reserve may reflect lower proximity to high schools, differences in school quality arising from funding gaps,[3] school organization or teacher characteristics, or differences in the characteristics of the students themselves.

Richards et al. (2008) estimate that approximately one in six students with Aboriginal identity in Canada attends a band-run school. Applying their estimation method to the same data, we find the proportion in B.C. is substantially lower. Data from the 2006 Census indicates that approximately 24% of 10 to 14 year-olds with Aboriginal identity in B.C. live on-reserve (Statistics Canada 2008). Data from the 2004 INAC Nominal Roll indicate that approximately 30% of on-reserve grade 1-8 students in B.C. attend a band-run school (Indian and Northern Affairs Canada 2005, Table 3.2). Taken together, these numbers imply that roughly 7.2 percent of grade 7 Aboriginal students in B.C. attend an on-reserve band-run school. This number is less than half of Richards et al.'s Canada-wide estimate, primarily because 70% of on-reserve students attend provincial schools in B.C. compared to 30% in Canada as a whole.

2.2 Access and funding within the provincial system

All students in B.C. are guaranteed placement in their neighborhood or "catchment area" public school. In addition, most public school districts offer magnet programs. French Immersion is by far the most popular magnet program in the province, enrolling about 5 percent of elementary school students. Prior to 2003, if a student wanted to register at a non-catchment public school other than through a magnet program, permission was required from both the principal of the catchment area school and the principal of the school of registration. In 2003, the Province instituted an official "open boundaries" policy that allows any student in B.C. to attend any public school if there are spaces available after local students have enrolled. It is not known whether this policy change has had a quantitatively important effect on cross-boundary

[3] Postl (2005) finds that the majority of band-run schools in B.C. are substantially under-funded relative to comparable provincial schools.

enrolments. In addition, approximately 10 percent of students in the province attend a private school.

The provincial Ministry of Education establishes curricula and provides operating and capital grants to the district school boards, who then allocate funds to individual public schools. District funding levels are based on a number of factors, including per-school allocations and per-student allocations, with supplementary funding based on the presence and number of students in several categories. In particular, districts receive supplemental funding for each Aboriginal student enrolled, for students with special educational needs, and for students who require English as a Second Language (ESL) services. Per student funding levels before and after 2002, when several major changes to the funding formula were introduced, are summarized in Table 1.

This provincial public school funding formula means that districts do not have to rely on local sources of revenue, ensuring that rich and poor districts receive similar funding. Indeed, because supplementary funding is targeted towards students with greater educational needs, districts with a greater number of disadvantaged students receive more funds than those with fewer disadvantaged students.

Private schools also receive per-student operating grants of 35-50% of the base public school rate, and are responsible for both teaching the provincial curriculum and meeting various provincial administrative requirements (British Columbia Ministry of Education 2005).

2.3 Data description

The administrative data used in this study are drawn from the Ministry's enrolment database and its Foundations Skills Assessment (FSA) exam database. Since the 1999/2000 school year, B.C. has administered the FSA tests in May of each year to students in grades 4 and 7 in all public and provincially funded private schools in British Columbia.[4] These exams are based on a variety of questions, both multiple-choice and open-ended, and are graded by accredited B.C. teachers. All students are expected to participate, with the exception of students in ESL programs who have not yet developed sufficient English skills to respond to the test, and some special needs

[4] FSA tests were also administered to grade 10 students between 2000 and 2004; these low-stakes assessments were replaced by high-stakes Provincial examinations beginning in 2005.

students. The FSA exams are relatively low-stakes for all parties. Students' scores do not contribute to their classroom grades and play no role in grade completion. The results do not affect school or district funding. However, school and district-level results are made public and are widely discussed within both the educational system and the news media. In particular, the Ministry of Education posts school-level results on its website, and a private research and advocacy organization produces a widely-publicized and much-discussed annual 'report card' that ranks all of the elementary schools in the province using a methodology based on FSA results (e.g. Cowley and Easton 2004).

Each B.C. student has a unique identification code, and we use an encrypted version of this code to link records across the enrolment and FSA exam databases, and to construct a longitudinal record for each student. Records in the enrolment database are based on Form 1701, the annual enrolment form collected for each student on September 30 of each year. These forms are used by the Ministry to determine school-level operational funding in accordance with the funding formulas described in Section 2.2. The enrolment record includes the student's current grade, school and district identifiers, year, gender, self-reported Aboriginal identity, enrolment in a language program (e.g. ESL, French Immersion, Francophone education), enrolment in a special needs program, and language spoken at home.[5] Records in the FSA exam database include the student's score on each exam, along with a flag indicating whether the student was excused from writing a given exam.[6]

Our data set covers all grade 4 and grade 7 students from the 1999-2000 through 2003-2004 school years. As a result, we observe the first three cohorts from the time they entered grade 4 in

---

[5] Because of confidentiality restrictions, our study is based on an extract from the original administrative data. The extract differs from the original data in the following ways: (1) enrolment records are provided only for students in grades 4 through 7; (2) student, school, and district identification codes are encrypted in such a manner as to allow for within-database linkage, but not linkage with external information; (3) language spoken at home is aggregated from the over 100 languages in the administrative data into English, Chinese (including Cantonese and Mandarin), Punjabi, and Other; and (4) both language spoken at home and Aboriginal status are provided based on the student's entire history rather than on the current year's self-report. In particular a student is categorized as Aboriginal if he/she ever self-reports as Aboriginal. A student is categorized as speaking English if he/she always self-reports as English, and is otherwise categorized by his/her most frequently reported home language other than English.

[6] Exam scores are calculated from item-level responses based on an item response theory (IRT) model constructed by the Ministry. The IRT scores are provided by the Ministry on a continuous scale with roughly zero mean and unit standard deviation. We normalize the scores in each year, grade, and subject to have exactly zero mean and unit standard deviation across the province.

1999, 2000 or 2001 through the end of their anticipated grade 7 year three years later. Using the unique student identifier, we link the records of students across multiple years to construct a panel of students who were in grade 7 between 2002 and 2004.[7]

Over 9% of the students in our data are reported by their parents or guardians as having Aboriginal identity. The extent to which this figure is an over- or under-estimate of the true proportion is unclear. On one hand, the availability of supplementary funding may lead schools to encourage parents and guardians to identify their children as Aboriginal. On the other hand, anecdotal reports suggest that some parents and guardians are reluctant to identify their children as Aboriginal within the school system because of concerns about stigma or discrimination. The proportion of students identifying as Aboriginal in our data is roughly comparable to Census-based figures; 8.2% of B.C. children aged 5-14 were identified as Aboriginal by their parents or guardians in the 2006 Census (Statistics Canada 2008).

3 The test score gap between Aboriginal and non-Aboriginal students

3.1 The simple test score gap

Table 2 presents our measures of academic achievement for Aboriginal and non-Aboriginal students. Each statistic is calculated using the full sample of students who have data on the exam(s) in question; for example, the mean grade 7 numeracy score is calculated using all students that took the grade 7 numeracy exam. The results (reported in appendix Table A1) are similar if one restricts the sample to students that took all four exams (i.e., both subjects in both grades 4 and 7).

As indicated in Table 2, Aboriginal students in grade 7 score more than 0.6 standard deviations on average below non-Aboriginal students on both exams. The results by quartile are similar: the achievement gap ranges from 0.51 to 0.77 standard deviations. Among students who wrote

[7] A minority of students who are observed in both grades 4 and 7 repeat grades, skip grades, or are out of province for one or more of the intervening years. We keep these students in our analysis whenever possible. If the student repeats either grade 4 or grade 7, we use the last year in grade 4 and the first year in grade 7.

the FSA numeracy test in both grades, the gap between the mean test scores of Aboriginal and non-Aboriginal students grew by an additional 0.05 standard deviations between grades 4 and 7, and the reading test score gap grew by 0.09 standard deviations. Again, the quartile results indicate that the gap in gains is not confined to any particular part of the distribution.

These gaps are comparable in magnitude to the black-white test score gap that has received a great deal of research attention in the U.S. (e.g., Card and Rothstein 2007; Cooley 2008; Fryer and Levitt 2004; Hanushek and Rivkin 2006). For example, the mean test score gap between blacks and whites on standardized numeracy tests in Texas elementary schools is about 0.76 standard deviations in grade 8, and grows by about 0.06 standard deviations between grades 3 and 8 (Hanushek and Rivkin, 2006, Table 3).

3.2 Accounting for exam participation

For various reasons, some students do not take the FSA exams. Any systematic differences between participants and nonparticipants may bias the results in Table 2. Exam participation may also be an outcome of direct interest. This section describes the participation patterns we observe in the data, and estimates bounds on the "true" achievement gap that are consistent with those patterns.

Table 3 reports exam participation rates for grade 7 students in B.C. from 2002 through 2004. Aboriginal students are more than twice as likely as non-Aboriginal students to miss a particular exam. About half of exam non-participants are excused from the exam. The other half simply do not take the exam, either because they are absent from school on exam day or because they do not respond to the exam. This low participation rate results in a high proportion of Aboriginal students with missing gain score data: about 29% on the numeracy exam and about 26% on the reading exam. While this is certainly a sufficiently high nonparticipation rate to be concerned about bias in our results, it should be noted that it is not out of line with the literature. For example, Hanushek et al. (2002) report exam participation rates for non-disabled and non-bilingual students in the well-known and heavily used Texas Schools Project data. For that relatively high-participation subgroup, they report grade 4 and grade 7 participation rates of 81.5% and 81.9% respectively.

Table 4 presents annual numeracy exam participation rates for 1999 through 2004. The results are similar for the reading exam. The table shows a clear trend towards both lower FSA participation of Aboriginal students and a higher proportion of Aboriginal students being excused from the exams.[8] Most of the growth in the proportion of students excused occurred between 1999 and 2001, but the downward trend in overall participation continued through the entire period of our data.

Table 5 shows how the characteristics of Aboriginal students differ between numeracy exam participants and nonparticipants. As might be expected, exam participation is not random. Students who failed to take the grade 7 exam without being excused were about three times as likely to have also missed the grade 4 exam as were grade 7 exam participants. Among those that were excused from the grade 7 exam, about 38% were also excused from the grade 4 exam, and another 16% simply failed to take it. Grade 7 unexcused nonparticipants that took the grade 4 exam scored 0.3 standard deviations less than grade 7 participants. The average score on the grade 4 exam among those that were subsequently excused from the grade 7 exam was 0.7 standard deviations less than grade 7 participants. Finally, students with disabilities account for a majority (about 62%) of excused absences from the exams, and a substantial proportion (about 28%) of unexcused absences.

The results in Tables 3-5 imply that nonparticipation rates are substantial, and that nonparticipants differ systematically from participants. This raises the issue of possible bias in the results in Table 2. In formal terms, Table 2 accurately describes the distribution of achievement conditional on Aboriginal identity and exam participation. However, our primary interest is in the distribution of achievement conditional on Aboriginal identity alone. In general, the distribution of interest is not identified in the absence of strong assumptions about nonparticipants but can often be bounded with weaker assumptions (Manski 1995).

---

[8] Tables 4 and 5 report participation rates for the numeracy exam. Participation rates for the reading exam are quite similar.

Table 6 reports bounds on mean and median achievement, under alternative assumptions about the achievement distributions of exam nonparticipants. The “worst-case” bounds (Manski 1995) impose no assumptions about the unobserved achievement of nonparticipants. The “increasing participation” bounds assume that participation rates are increasing in achievement, and the average achievement of nonparticipants is bounded from below at -2. The first of these assumptions is consistent with the characteristics of nonparticipants seen in Table 5, and less than one percent of exam participants score lower than 2 standard deviations below average on the FSA exam. The “exogenous participation” bounds are calculated under the conventional but implausible assumption that there is no systematic achievement difference between participants and nonparticipants once we condition on Aboriginal identity. In this case, the bounds collapse to the original point estimate.

The worst-case bounds in Table 6 are uninformative for the mean in the absence of bounds on the support of the achievement distribution, but they are quite informative for the median. For example, the median numeracy achievement of non-Aboriginal students is at least 0.18 standard deviations higher than that of Aboriginal students, and may be more than 1.1 standard deviations higher. The increasing participation bounds are informative for both the mean and median, generally implying a gap ranging from slightly less than a half standard deviation to slightly less than a full standard deviation. Comparing these bounds with the point estimate under the assumption of exogenous participation, we find that the gap among participants is not necessarily a lower bound on the true gap. To summarize these results, we find that the finding in Table 2 of a sizeable gap in test scores persists under substantially weaker assumptions about participation.

3.3 Accounting for student and school-level factors

Table 7 provides some insight into the student-level factors underlying the relatively low participation rates and test scores of Aboriginal students. The incidence of assessed disabilities is two and a half times as high in the Aboriginal population as in the non-Aboriginal population, and Aboriginal students are only one-third as likely as non-Aboriginal students to be assessed as gifted. Aboriginal students are overrepresented in every category of disability, but most dramatically in the severe behavioural disorder category. Almost 7% of Aboriginal students are found to have a moderate or severe behavioural disorder, compared to fewer than 2% of non-Aboriginal students.

Table 8 presents results from OLS regressions where the dependent variable is the level of the individual student's test score on the grade 7 numeracy or reading exam. [9] The first column uses only Aboriginal identity as an explanatory variable, and thus reproduces the difference in mean test scores of Aboriginal and non-Aboriginal students. The second column includes controls for gender and whether a student has a disability. The regression coefficients in column 2 show that, as expected, students with disabilities have much lower achievement levels than those who do not. Given their high rate of disabilities, developing effective programs for these special populations is of particular importance for Aboriginal students. However, in spite of these results, and in spite of the higher disability rates among Aboriginal children, the Aboriginal test score gap remains at almost 0.6 standard deviations on both tests when we condition on identified disability. This analysis suggests that improving outcomes for students in special education would have a limited effect on the overall Aboriginal/non-Aboriginal test score gap.

The third column of Table 8 adds school fixed effects to account for the possible role of school quality in generating the test score gap. The coefficient on Aboriginal identity declines by 40% for numeracy and by 33% for reading, but remains negative, statistically significant, and quite large. In other words, the test score gap has both a substantial within-school component (Aboriginal students on average have lower scores than non-Aboriginal students attending the same school) and a substantial between-school component (Aboriginal students on average attend schools with lower average achievement). Section 4.2 further explores quantifying and interpreting this decomposition into within-school and between-school components.

The fourth column of Table 8 adds interactions between Aboriginal identity and other individual characteristics. The results show that the male-female gap in test scores is smaller in absolute value for Aboriginal students. The positive coefficients on the interaction between Aboriginal

[9] The regressions in Table 8 are estimated using data on all students that took the grade 7 exam, while the regressions in Table 9 are estimated using only those students who also took the grade 4 exam. The results of estimating the Table 8 regressions using the Table 9 sample (reported in Appendix Table A2) are similar.

identity and learning/behavioral disabilities imply that the gap between Aboriginal and non-Aboriginal students is substantially smaller (but still present) among students with these disabilities. They also imply that the gap between disabled and non-disabled students is smaller for Aboriginal students than for non-Aboriginal students.

Table 9 presents estimates from a similar set of specifications where the dependent variable is the change in the student's reading or numeracy test score between grades 4 and 7. The results show that about 6% of the test score gap in numeracy in grade 7 and 14% of the test score gap in reading emerges after grade 4 (computed by dividing the coefficient on the Aboriginal dummy from column 1 in Table 9 by the same coefficient from column 1 in Table 8). Although policies that focus attention on the years before grade 4 have significant potential to improve later outcomes, Aboriginal children do continue to fall further behind on average between grades 4 and 7. When we include the additional control variables in this value-added specification in column 2, we again find that disabled students warrant particular policy attention. Conditional on these characteristics, Aboriginal students continue to fall behind their non-Aboriginal schoolmates at a rate that is of considerable policy significance. Adding fixed effects, as in column 3, reduces the coefficient on Aboriginal identity by about 50% for numeracy and 33% for reading. Like the gap in test score levels, the gap in gains has substantial within-school and between-school components. Finally, column 4 includes interactions between Aboriginal identity and other individual characteristics. The results here show an interesting pattern by gender: boys account for almost the entire gap in numeracy gains, while girls account for almost the entire gap in reading gains.

4 Sorting across schools

4.1 How much sorting is there?

Aboriginal and non-Aboriginal students are likely to exhibit different school attendance patterns for several reasons. First, Aboriginal and non-Aboriginal students tend to live in different communities. Aboriginal students in Canada are disproportionately located in small rural communities and a handful of urban centres (Statistics Canada 2008). Second, differential patterns of attendance at magnet and private schools also contribute to the overall pattern of sorting. Non-Aboriginal students are almost twice as likely to attend private schools as Aboriginal students are: 10.6% of non-Aboriginal students in our data attend private schools, compared to 5.7% of Aboriginal students. Another 6.2% of non-Aboriginal students are enrolled in French Immersion programs, compared to 2.4% of Aboriginal students.

Figure 1 shows the frequency distribution of the proportion of Aboriginal students among the grade 7 cohorts at all B.C. public and private schools in 2002, 2003 and 2004. In over 22% of cases, no Aboriginal students were enrolled in grade 7 within a school and year. Private schools and magnet schools play an important role in the sorting process: almost two-thirds of the school/years that had no Aboriginal students were associated with either private schools or French Immersion magnet programs.[10] In the modal grade 7 school/year in our sample, at least one student and fewer than 10% of students are Aboriginal; over 42% of all school/years fall into this category. In another 17% of school/years, between 10% and 20% of students are Aboriginal. At the other extreme, almost 3% of school/years include no non-Aboriginal students. The overall picture that emerges is one in which over 25% of the grade 7 school/years between 2002 and 2004 are fully segregated, and Aboriginal students are dispersed widely across schools that enroll both Aboriginal and non-Aboriginal students.

Figures 2 and 3 provide frequencies of the percentage of a student's peers who are Aboriginal, for Aboriginal and non-Aboriginal students respectively. Figure 2 shows that almost 5% of the Aboriginal students in our sample have no Aboriginal same-grade peers. More than 27% of Aboriginal students have more than zero but fewer than 10% Aboriginal peers. Another 23% of Aboriginal students attend schools where between 10% and 20% of their same-grade peers are Aboriginal. A substantial fraction of Aboriginal students attend schools in which Aboriginal students are more heavily concentrated, and over 5% of Aboriginal students have no non-Aboriginal peers. Figure 3 shows that, in contrast, almost 75% of non-Aboriginal students attend schools where fewer than 10% of same-grade peers are Aboriginal. The proportion of non-Aboriginal students who attend schools in which Aboriginal students are the majority is negligible.

---

[10] In schools that house both regular and French Immersion programs, we define the French Immersion program as a distinct "school" in all of our analysis.

The extent to which Aboriginal and non-Aboriginal students are distributed differently across schools can also be summarized with a standard dissimilarity index (Duncan and Duncan 1955). According to this measure, almost half of Aboriginal students would have to change schools in order to achieve an equal distribution of the two groups across schools.[11]

4.2 The implications of sorting for within/between-school decompositions

The fixed effects results in Section 3.3 imply that the outcome gap has both within-school and between-school components: on average, Aboriginal students attend schools with worse average outcomes for all students, and they have worse average outcomes than non-Aboriginal students attending the same school. These stylized facts are similar to those documented for black and white students in the U.S. (e.g. Fryer and Levitt 2004; Hanushek and Rivkin 2006). We follow that literature and provide a quantitative decomposition of the test score gap into within-school and between-school components.

The purpose of this kind of decomposition is to get an idea of the potential role of school quality in creating the outcome gap, though any such interpretation should be made with a clear eye on its limitations. If students are randomly assigned to schools conditional on Aboriginal identity, and the treatment effect of enrollment in a given school is constant across students (in particular, the school's effect does not differ for Aboriginal and non-Aboriginal students), then the between-school gap may be an accurate measure of the contribution of variations in school environment to the overall gap.[12] If instead of random assignment there is positive sorting on unobserved outcome-relevant characteristics such as socioeconomic status, some portion of the school fixed effects will actually represent these characteristics. In that case, the between-school gap can be considered an upper bound on the portion of the gap that is caused by variations in school

---

[11] The dissimilarity index is calculated as $D = \frac{1}{2}\sum_{s=1}^{S}\left|\frac{n_s^A}{n^A} - \frac{n_s^{NA}}{n^{NA}}\right|$, where n_s^A is the number of Aboriginal students attending school s; n^A is the total number of Aboriginal students; n_s^{NA} is the number of non-Aboriginal students attending school s; and n^{NA} is the total number of non-Aboriginal students. The index values for our data in all three years are very similar; the 2004 value is 48.5.

[12] The school fixed effects capture the influence of peers as well as teachers and other school characteristics, and so even under these strong assumptions are not invariant to a reallocation of students across schools.

quality. If instead of a constant treatment effect, each school has differential effects on Aboriginal and non-Aboriginal students, then things are somewhat more complex. For example, if each school provides lower average quality instruction to Aboriginal students than to non-Aboriginal students, this will appear in the within-school component of both decompositions even though it is driven entirely by school quality. If Aboriginal students tend to attend schools that are more effective for Aboriginal students relative to non-Aboriginal students than the average school in the province, then the between-school component of the decomposition still provides an upper bound on the potential gains from school changes. Finally, we note that school fixed effects include the quality of peers. If peer effects are substantial, a reallocation of students between "low-quality" and "high-quality" schools may change the relative quality of the schools. In that case, decompositions may overstate or understate the potential aggregate gains from moving Aboriginal students to high-quality schools even if they correctly measure the gains for an individual Aboriginal student. This issue is further discussed in Section 5.2.

The literature uses two distinct decomposition approaches. Fryer and Levitt (2004, Table 6) use the absolute value of the coefficient on Aboriginal identity from the regression with school fixed effects but no additional controls ($\hat{\delta}_{FE}$) as a measure of the within-school outcome gap. The between-school gap is the difference in average school fixed effects from that same regression:

$$\underbrace{\bar{y}^{NA} - \bar{y}^{A}}_{Total\ gap} = \underbrace{\left[\sum_{s=1}^{S}\left(\frac{n_s^{NA}}{n^{NA}} - \frac{n_s^{A}}{n^{A}}\right)\left(\bar{y}_s - \hat{\delta}_{FE}\frac{n_s^{NA}}{n_s}\right)\right]}_{FL\ Between-school\ gap} + \underbrace{\hat{\delta}_{FE}}_{FL\ Within-school\ gap} \qquad (1)$$

where $\bar{y}$ is an average test score, n is a student count, the superscript indicates aboriginal status (NA = non-Aboriginal, A = Aboriginal, no superscript = all students) and the subscript indicates the school (no subscript = all schools). Hanushek and Rivkin (2006, equation 1) argue that Fryer and Levitt's approach potentially understates the contribution of schools and propose an alternative decomposition that takes the form:[13]

---

[13] The difference between the two decompositions can be explained with an example adapted from Hanushek and Rivkin. Suppose there are only 3 schools: one in which all students are Aboriginal and the average score is *a*, one in which all students are non-Aboriginal and the average score is *b*, and one that is integrated with an average score of *c* for Aboriginal students and *d* for non-Aboriginal students. In addition, assume that the segregated schools account for an arbitrarily large portion of the students. In the Fryer and Levitt decomposition, the within-school gap would be *(d-c)*, and the between-school gap would be *(b-a)-(d-c)*. In the Hanushek-Rivkin decomposition, the within-school gap would be zero and the between-school gap would be *(b-a)*. As one might expect from this example, the

$$\underbrace{\bar{y}^{NA} - \bar{y}^{A}}_{Total\ gap} = \underbrace{\left[\sum_{s=1}^{S}\left(\frac{n_s^{NA}}{n^{NA}} - \frac{n_s^{A}}{n^{A}}\right)\bar{y}_s\right]}_{HR\ Between-school\ gap} + \underbrace{\left[\left(\frac{1}{n^{NA}} + \frac{1}{n^{A}}\right)\sum_{s=1}^{S}(\bar{y}_s^{NA} - \bar{y}_s^{A})\,\frac{n_s^{NA}n_s^{A}}{n_s}\right]}_{HR\ Within-school\ gap} \quad (2)$$

$$= \underbrace{\left[\sum_{s=1}^{S}\left(\frac{n_s^{NA}}{n^{NA}} - \frac{n_s^{A}}{n^{A}}\right)\bar{y}_s\right]}_{HR\ Between-school\ gap} + \underbrace{\left(1 - \sum_{s=1}^{S}\left(\frac{n_s^{NA}}{n^{NA}} - \frac{n_s^{A}}{n^{A}}\right)\frac{n_s^{NA}}{n_s}\right)\hat{\delta}_{FE}}_{HR\ Within-school\ gap}$$

Table 10 reports both decompositions. According to both methods, the between-school component of the gap in grade 7 test scores is substantial; ranging from 34% (Fryer-Levitt) to 45% (Hanushek-Rivkin) in reading, and from 40% to 49% in numeracy. The between-school component of the gap in reading test scores gains is slightly larger than in reading test score levels, ranging from 37% to 47%, and is substantially larger in numeracy, ranging from 59% to 63%. While substantial, between-school factors appear to be somewhat less important in B.C. than in Texas, where Hanushek and Rivkin (2006) report that over 75% of the growth between grades 3 and 8 in the black/white test score gap is accounted for by between-school factors.

5 How important are peers?

Given the substantial amount of sorting of Aboriginal and non-Aboriginal students across schools, the differences in their achievement levels and growth, and the sizable between-school component of the overall mean test score gap, the hypothesis that differences in school quality contribute to the relatively low achievement levels of Aboriginal students warrants serious consideration. Hanushek and Rivkin (2006) and Hanushek, Kain and Rivkin (2009) find that observable school-level factors including teacher experience, student turnover and the racial composition of the student body explain a significant proportion of the black/white achievement gap in Texas. Our data do not include measures of school inputs such as teacher salaries, accreditation and experience or class size. In any case, the provincial funding formula described earlier implies that variation in the quality of teaching inputs across schools is likely to be considerably smaller in B.C. than in U.S. jurisdictions that rely on local taxation to support schools. However, if peer effects are important, differences in peer group composition may be

---

Hanushek-Rivkin decomposition generally produces a larger between-school component than does the Fryer-Levitt decomposition.

an important dimension along which Aboriginal and non-Aboriginal students experience differences in average school quality.

The term “peer effects” is used in this paper to incorporate a wide variety of direct and indirect effects of the composition of a student’s peer group on his or her educational outcomes. Some peer effects operate within the classroom and on a daily basis, while other peer effects operate at the school level and/or over a longer term. Within the classroom, a low-achieving peer group may reinforce negative social norms, or may provide fewer opportunities for students to learn from other students. Families with limited resources may be unable to supply classroom public goods such as volunteering, fundraising, and monitoring. Students with behavioral disorders or learning disabilities may take instruction time or attention away from classmates. Cultural factors may also affect classroom dynamics. For example, Aboriginal children with more Aboriginal classmates may experience less racism and may find more support for a positive cultural identity, and their parents may be more inclined to participate in the school. In addition to these direct effects, peers may also affect the learning environment indirectly by affecting resource allocations. For example, the provincial funding formulas described in Table 1 imply that additional resources are likely to follow both Aboriginal students and students with special needs. Principals may also adjust teacher assignments, class sizes, or other resources in response to year-to-year variation in student needs.

Most of the daily classroom-level mechanisms of peer influence on achievement also operate at the grade level, and at the school level. In addition, a school’s composition may have substantial long-term effects on teacher quality. Hanushek, Kain, and Rivkin (2004) find that teachers in Texas have a much higher rate of exit from schools with a greater proportion of disadvantaged, minority, or low-achieving students. Scafidi et al. (2007) find that similar exit patterns in Georgia are driven primarily by teachers exiting schools with high minority shares. If inexperienced teachers tend to be less effective, or if effective teachers are more mobile than are ineffective teachers, then schools with a higher proportion of minority students will tend to have less effective teachers. By analogy, these U.S.-based results suggest that schools with a larger proportion of Aboriginal students may face substantial difficulty in hiring and retaining effective teachers.

These examples suggest that the peer environment of Aboriginal students can either help or hinder their achievement, making the net effect an unsettled empirical question. A growing body of evidence with respect to the role of racial segregation and peer effects in the U.S. has produced mixed results (e.g. Rivkin and Welch 2006, Card and Rothstein 2007). In the Canadian context, Friesen and Krauth (2008) find significant peer effects associated with some home language groups on the test scores of non-Aboriginal students in data from British Columbia drawn from the same administrative file used in this paper.

Richards et al. (2008) use cross-sectional data based on the same data source to estimate the relationship between the school-level proportion of Aboriginal students who "meet or exceed expectations" (the meets/exceeds expectations ratio or MER) and two variables intended to capture what they call "in-school dynamics" or peer effects. Their school-level regressions include measures of student socioeconomic characteristics and several variables that are measured at the school district level. They find a positive relationship between the school-level Aboriginal MER and the MER of non-Aboriginal students in the same school, and a negative relationship between the school-level Aboriginal MER and the number of Aboriginal students who write the exam. They interpret the first relationship as evidence of positive spillovers from high-achieving non-Aboriginal students to their Aboriginal schoolmates, and the second relationship as evidence of negative peer effects among Aboriginal students.

However, these causal interpretations are questionable. Manski (1993, 1995) demonstrates that peer effects are in general not identified from cross-sectional data whenever the assignment of individuals to peer groups is nonrandom. The assignment of a student to a school is a clear example of nonrandom selection. Families make choices about residential location, private schooling, and enrollment in special programs like French Immersion, and these choices determine where a child attends school. The family's resources and preferences will affect these choices, and will affect the allocation of parental time. If preferences or resources vary systematically by background characteristics, the background characteristics of a student's peer group will be correlated with both school quality and the student's own private resources.

In this case, the measured cross-sectional relationship between the MER of Aboriginal and non-Aboriginal students in the same school will provide a biased measure of any true peer effects if there is unobserved variation in school inputs that affects non-Aboriginal and Aboriginal students in the same direction, or if the Aboriginal students who choose to attend schools where non-Aboriginal students perform better than expected are themselves above-average. Both of these scenarios seem quite plausible. The measured cross-sectional relationship between the school-level Aboriginal MER and the number of Aboriginal students with test scores will provide a biased estimate of peer effects if Aboriginal students who are more disadvantaged tend to be concentrated in a subset of schools, or if schools with more Aboriginal students have fewer resources or less experienced teachers.[14] Our approach to estimating Aboriginal peer effects, described next, addresses these potentially confounding issues.

5.1 Model specification and research design

In order to distinguish between peer effects and unobserved factors that could confound our estimates, we use a now-standard method for estimating peer effects in education that uses individual student-level panel data from multiple cohorts of students within each school,[15] and exploits the small but plausibly random year-to-year variation in peer group composition within a school to consistently estimate grade- level peer effects, while allowing for systematic cross-school variation in school or student quality. This approach has been used by a number of authors in varying contexts (e.g. Hoxby 2000; Hanushek, Kain and Rivkin 2009; Cooley 2008; Lavy, Paserman and Schlosser 2007), and provides relatively clean identification but with some important limitations. First, it captures only within-grade peer effects, including within-classroom effects. Peer effects operating at the school or neighborhood level will be included in the school fixed effect. Most of the mechanisms underlying school-level peer effects also appear at the grade level, so it may be reasonable to interpret the sign and relative magnitude of grade-

---

[14] A further problem arises because the authors do not include Aboriginal students who do not write the FSA exams in their count of Aboriginal peers.

[15] Like both Hoxby and Hanushek et al., we measure peer composition at the grade rather than the classroom level and therefore avoid selection effects associated with classroom assignment. Betts and Zau (2005) are able to distinguish between classroom and grade-level peer effects in their administrative data set from San Diego. Their results indicate that most of the effect of peers' achievement on individual achievement is related to classroom peers in mathematics, and to both classroom and grade level peers in reading. They find that including only grade level peer effects results in somewhat smaller coefficient estimates, but does not change the overall pattern of the results.

level effects as evidence on the sign and relative magnitude of peer effects at a higher level. However, such an interpretation relies heavily on extrapolation. The second limitation arises because within-school variation in peer group composition is relatively small. Using the measured effect of these small within-school changes in composition to infer the effect of larger between-school differences also relies heavily on extrapolation, and may miss important nonlinearities.

The model is constructed as follows. Students are indexed by $i=1,2,\ldots,n$; schools by $s=1,2,\ldots,S$; grades by $g=4,7$; and time by $t=1,2,\ldots,T$. FSA exam subjects (i.e. reading and numeracy) are indexed by $j=1,2$. Let $y^j_{i,g}$ be the score of student i on exam j in grade g. Let $t(i,g)$ be the school year in which the student takes grade g, and let $s(i,g)$ be the school student i attends in grade g. Let $X_{i,g}$ be a vector of student i's individual background characteristics in grade g, and let the vector $\bar{X}_{i,g}$ be the average value of X among student 'is same-grade schoolmates in grade g. Our value-added regression model takes the form:

$$y^j_{i,7} - y^j_{i,4} = \beta^j X_{i,7} + \lambda^j \bar{X}_{i,7} + \delta^j_{t(i,7)} + a^j_{s(i,7)} + v^j_{s(i,7),t(i,7)} + u^j_{i,7} \qquad (3)$$

$$E\left(v^j_{s(i,7),t(i,7)} + u^j_{i,7} \middle| X_{i,7}, \bar{X}_{i,7}, \delta^j_{t(i,7)}, a^j_{s(i,7)}\right) = 0 \qquad (4)$$

where β^j and λ^j are vectors of parameters to be estimated, $\delta^j_{t(i,7)}$ is an unobserved year-specific fixed effect, $a^j_{s(i,7)}$ is an unobserved school-specific fixed effect, $v^j_{s(i,7),t(i,7)}$ is an unobserved school-and-year-specific effect and $u^j_{i,7}$ is an unobserved individual-specific effect The content of our identifying assumption is similar to that in the related literature; while the overall composition of a school may be systematically related to unobserved school and student characteristics, the small cohort-to-cohort fluctuations in composition within a school may be considered essentially random and thus unrelated to cohort-to-cohort fluctuations in other unobserved factors. We allow for within-group common shocks like, for example, an instructor being replaced with a lower-skilled substitute while on parental leave, provided these shocks are unrelated (in conditional mean) to the observed composition of the group.

Equation (3) takes the same form as the simple value-added (SVA) model commonly used to estimate education production functions in the education literature. An alternative, sometimes called the modified value-added (MVA) model, adds the lagged test score as a control variable. Both models are often interpreted as reduced-form estimating equations for a structural model in which the current test score is a function of cumulative inputs to the production function for student knowledge. Todd and Wolpin (2003) show that this interpretation requires two strong assumptions. First, the earlier test score must be a sufficient statistic for all relevant prior inputs, including initial ability. Second, all current inputs that are correlated with the earlier test score must be observed. This rules out, for example, any increased attention by parents and schools in response to a student's disappointing performance on the first exam. The SVA model further requires that these prior inputs should have the same effect on current performance as on past performance. While the MVA model thus has the advantage of being more general, we use the SVA model because the MVA model risks substantial bias from measurement error in the earlier test score. In either case, Todd and Wolpin's critique of value-added models suggests that our results are best interpreted as measuring a reduced form "policy" effect rather than structural parameters of the production function. [16]

Finally, we note that nonparticipation has the potential to bias our results unless:

$$E\left(v^{j}_{s(i,7),t(i,7)} + u^{j}_{i,7} \middle| X_{i,7}, \bar{X}_{i,7}, \delta^{j}_{t(i,7)}, \alpha^{j}_{s(i,7)}, p^{j}_{i,7} = 1\right) = 0 \quad (5)$$

where p is an indicator of exam participation. That is, exam participation may differ systematically by school, year, and observed characteristics, but cannot differ by any unobserved individual-specific or school-year-specific factors that affect value added outcomes. To the extent that students with low value-added achievement are overrepresented among nonparticipants, our estimated effects may be understated.

[16] A further complication in interpreting our regression coefficients as parameters of a cumulative-input education production function is introduced by the fact that there is a three-year gap between exams. Our main regressions only include measures of grade 7 inputs, including both peer characteristics and the school fixed effect. With unlimited data it would be preferable to include grade 5 and 6 inputs as well, including grade-specific school fixed effects. As such an approach would rapidly exhaust degrees of freedom in our regressions, we prefer to estimate models with grade 7 inputs only. These results should be interpreted with the caveat that grade 7 peer characteristics are also acting as a proxy for grade 5 and 6 peer characteristics. This issue would still be present (though to a substantially lesser degree) in data with annual testing: students change schools during the year, and the peer group measured on a particular day during the year is used as a proxy for the peer group during the year as a whole.

5.2 Results

The population in our regression analysis is B.C. public and private school students who attended grade 7 between 2002/2003 and 2004/2005, and who were enrolled in grade 4 in B.C. in 1999/2000 or later. All specifications are estimated from the population of Aboriginal students for whom the relevant outcome is observed, while the school-grade compositional variables are based on the entire population of enrolled students, including both non-Aboriginal students and students who do not take the exam. The individual-level control variables include gender and current special needs category, if applicable, and the peer measures include the proportion of same-grade peers who are Aboriginal, male, and classified as having various types of disabilities. Table 11 provides means of the peer variables for Aboriginal and non-Aboriginal students; means for the other variables are reported in Table 7.

Our regression results are reported in Table 12. School and year fixed effects are included in all specifications, and estimated standard errors are robust to clustering at the school-year level. Peer group composition is being reported in decimal rather than percentage units, so each coefficient can be interpreted as the exam score increase (in standard deviations) associated with the percentage of peers in a given category increasing from 0% to 100%. The specification in column (1) includes an individual control for gender along with the percent male and percent Aboriginal. The specification in column (2) adds controls for own and peer disability status, while column (3) distinguishes between learning/behavioural disability and other disability.[17]

The coefficients for the individual characteristics differ in some cases from the estimates in Table 9. In particular, the coefficient on behavioural disabilities is smaller and statistically insignificant when peer characteristics are included in the model and the sample includes Aboriginal students only. On the other hand, having a disability in the "other" category appears to have a larger effect on reading scores than would be suggested by the results in Table 9.

---

[17] This disaggregation of special needs into three categories follows work by Hanushek, Kain and Rivkin (2002) and Friesen, Hickey and Krauth (2009). Learning and behavioural disabilities together account for about 70% of disabled students in B.C.

Turning to the peer effects estimates, our results are consistent with those found elsewhere in the literature that male peers are associated with lower test score gains (e.g. Hoxby 2000, Lavy and Schlosser 2007). This result is statistically significant in the case of the numeracy exam. Interestingly, the effect of Aboriginal peers is positive in all specifications, and it is statistically significant in our base specification for the numeracy exam. Peers with learning or behaviour disabilities are associated with lower test score growth in numeracy, although these estimates again are statistically insignificant. Moreover, they are quite small in magnitude given the range of variation in our data. Peers who have other disabilities have a positive influence on Aboriginal students' test score gains. This result is consistent with other research for non-Aboriginal students in B.C. that finds a positive effect of students with "other" disabilities on the test scores of disabled non-Aboriginal students (Friesen, Hickey and Krauth 2009), and could be the result of the extensive resources directed towards some of these students.

To get an idea of the magnitude of these peer effect estimates, Table 12 also includes an estimate of the "Total peer effect." The total peer effect is defined here as the change in predicted outcome implied by a change in peer group from one identical to the average peer group of non-Aboriginal students to one identical to the average peer group of Aboriginal students, i.e.:

$$TPE = \hat{\lambda}(\bar{X}^{A} - \bar{X}^{NA}) \qquad (6)$$

where $\hat{\lambda}$ is the estimated vector of peer effect coefficients, $\bar{X}^{A}$ is the vector of average peer group composition for Aboriginal students and $\bar{X}^{NA}$ is the corresponding vector for non-Aboriginal students. The reported standard errors are estimated treating $(\bar{X}^{A} - \bar{X}^{NA})$ as fixed and using the cluster-robust covariance matrix estimator for $\hat{\lambda}$.

As the table shows, the total peer effect is consistently positive but statistically significant: the typical peer group of an Aboriginal student is estimated to increase his or her numeracy test score by 4.4 to 4.7 percent of a standard deviation, and to increase his or her reading score by approximately 1.7 percent of a standard deviation. These quantities are small relative to the gap in test score levels (60% of a standard deviation for both subjects), but large in comparison to the gap in test score gains (5% of a standard deviation for numeracy and 9% of a standard deviation for reading). One limitation of this measure of the total peer effect is that it only includes those avenues of peer influence that are detectable through year-to-year variation within a school.

Long-run effects of school composition on teacher quality, social norms, and resource availability may not appear in this calculation.

Finally, these results have some implication for the achievement gap decompositions in Section 4.2. These decompositions aim to characterize the proportion of the achievement gap that might be attributable to differences in school quality. A nonzero peer effect complicates this interpretation, because school quality as measured by these decompositions includes peer quality and therefore is not invariant to who attends the school. For example, suppose that the only difference between "good" school A and "bad" school B is peer quality. In that case, an individual student would benefit from a move from school B to school A, but switching all students between the two schools would leave all student outcomes unchanged. The case here is actually somewhat different: because the effect of Aboriginal peers is positive, the difference in school quality excluding peer quality is actually larger than the difference including peer quality.

6 Conclusion

Our estimates show that about half of the growth in the test score gap between Aboriginal and non-Aboriginal students between grades 4 and 7 can be accounted for by between-school factors, so the potential role of differences in school characteristics in explaining the overall achievement gap is substantial. However, the funding formula used in B.C. directs greater resources into school districts with greater numbers of Aboriginal students and students with special needs, so it seems unlikely that the achievement gap between Aboriginal and non-Aboriginal students is explained by a relative lack of financial resources in schools that Aboriginal students attend. It is possible that these schools are less successful at attracting skilled teachers; unfortunately, our data do not allow us to explore this hypothesis. Aboriginal and non-Aboriginal students are sorted across schools so that the average Aboriginal student has a substantially higher proportion of Aboriginal peers and peers with disabilities, and we implement a methodology that allows us to measure plausibly causal effects of small variations in peer composition on student test score growth.

Our econometric evidence provides little support for the hypothesis that peer composition contributes to the between-school component of the growth in the test score gap between

Aboriginal and non-Aboriginal students. If anything, Aboriginal students may benefit from attending school with higher concentrations of Aboriginal students and higher concentrations of students with some disabilities, perhaps because these students bring additional funds. We find weak evidence that Aboriginal students' disproportionate exposure to students with learning disabilities and behavioural disorders may have a moderate adverse impact on their achievement. These results are consistent with the results of other research on peer effects associated with disabled students (Hanushek et al. 2002, Friesen, Hickey and Krauth 2009). Given the absence of evidence that Aboriginal students' peers have a substantial influence on their academic performance, it is tempting to conclude that school choice policies, such as voucher systems, could not contribute much to the academic achievement of Aboriginal students. However, it is important to bear in mind that our methodology uses small year-to-year changes in peer composition within a school to identify peer effects, and the results may not generalize to larger differences between schools in peer composition, particularly if there are substantial nonlinearities either in peer effects or in the long-run relationship between school composition and school operations.

References

British Columbia Ministry of Education, 2002. *2002/03 Operating Grants Manual to British Columbia School Boards.* Victoria: Government of British Columbia. Available at http://www.bced.gov.bc.ca/k12funding/funding/02-03/estimates/operating-grants-manual.pdf

________, 2005. *Overview of Independent Schools in British Columbia*. Victoria: Government of British Columbia.

Betts, Julian R. and Andrew Zau, 2005. Peer groups and academic achievement: Panel evidence from administrative data. Working paper, University of California, San Diego.

Card, David, 1999. The causal effect of education on earnings, in Orley Ashenfelter and David Card eds., *Handbook of Labor Economics, Volume 3A.* Amsterdam: Elsevier North-Holland.

Card, David and Jesse Rothstein, 2007. Racial segregation and the black-white test score gap, *Journal of Public Economics* 91(11-12): 2158-2184.

Cooley, Jane, 2008. Desegregation and the achievement gap: Do diverse peers help? Working paper, University of Wisconsin-Madison.

Cowley, Peter and Steven Easton, 2004. *Report Card on British Columbia Elementary Schools*. Vancouver: Fraser Institute.

Dee, Thomas S., 2004. Are there civic returns to education? *Journal of Public Economics* 88: 1696-1712.

Duncan, Otis D. and Beverly Duncan, 1955. A methodological analysis of segregation indexes, *American Sociological Review* 20: 210–217.

Friesen, Jane and Brian Krauth, 2008. Enclaves, peer effects and student learning outcomes in British Columbia, Working Paper, Simon Fraser University.

Friesen, Jane, Ross Hickey and Brian Krauth, 2009. Disabled peers and academic achievement, *Education Finance and Policy*, forthcoming.

Fryer, Roland G. and Steven D. Levitt, 2004. Understanding the black-white test score gap in the first two years of school, *The Review of Economics and Statistics* 86: 447-464.

Greenwood, Daphne T., 1997. New developments in the intergenerational impacts of education, *International Journal of Education Research* 27: 503-512.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin, 2002. Inferring program effects for special populations: Does special education raise achievement for students with disabilities? *Review of Economics and Statistics* 84: 584-599.

____, 2004. Why public schools lose teachers, *Journal of Human Resources* 39: 326-354.

____, 2009. New evidence about Brown v. Board of Education: The complex effects of school racial composition on achievement. *Journal of Labor Economics* 27: 349-385.

Hanushek, Eric A. and Steven G. Rivkin, 2006. School quality and the black-white achievement gap, NBER Working Paper 12651, National Bureau of Economic Research.

Health Canada, 2009. *A Statistical Profile on the Health of First Nations in Canada: Determinants of Health 1999-2003*. Ottawa: Health Canada. Available at http://www.hc-sc.gc.ca/fniah-spnia/pubs/aborig-autoch/index-eng.php.

Hoxby, Caroline M., 2000. Peer effects in the classroom: Learning from gender and race variation. NBER Working Paper 7867, National Bureau of Economic Research.

Indian and Northern Affairs Canada, 2005. *Basic Departmental Data 2004*. Ottawa: Indian and Northern Affairs Canada. Available at http://www.ainc-inac.gc.ca/ai/rs/pubs/sts/bdd/bdd-eng.asp.

Kenkel, Donald S., 1991. Health behavior, health knowledge and schooling, *Journal of Political Economy* 99: 287-305.

Lavy, Victor and Analia Schlosser, 2007. Mechanisms and impacts of gender peer effects at school, NBER Working Paper 13292, National Bureau of Economic Research.

Lavy, Victor, M. Daniele Paserman and Analia Schlosser, 2007. Inside the black box of peer ability effects: Evidence from variation in the proportion of low achievers in the classroom, NBER Working Paper No. 14415, National Bureau of Economic Research.

Lochner, Lance and Enrico Moretti, 2004. The effect of education on crime: Evidence from prison inmates, arrests and self-reports, *American Economic Review* 94: 155-189.

Manski, Charles F., 1993. Identification of endogenous social effects: The reflection problem, *Review of Economic Studies* 60: 531-542.

________, 1995. *Identification Problems in the Social Sciences*. Cambridge MA: Harvard University Press.

Mendelson, Michael, 2006. *Aboriginal Peoples and Post-Secondary Education in Canada*. Ottawa: Caledon Institute of Social Policy.

Postl, Bill, 2005. *British Columbia First Nations Schools Funding Analysis: 2003/04 School Year*. Report to First Nations Education Steering Committee, available at http://www.fnesc.ca/Attachments/Publications/PDF%27s/pdf/Bill%20Postl%20Report/0304%20bc%20revised.doc.

Richards, John and Aidan Vining, 2004. *Aboriginal Off-reserve Education: Time for Action*. Toronto: C.D. Howe Institute.

John Richards, Jennifer Hove, and Kemi Afolabi, 2008. *Understanding the Aboriginal/Non-Aboriginal Gap in Student Performance: Lessons from British Columbia*. Toronto: C.D. Howe Institute. Available at http://www.cdhowe.org/pdf/commentary_276.pdf.

Rivkin, Steven and Finis Welch, 2006. Has school desegregation improved academic and economic outcomes for blacks?, in Eric Hanushek and Finis Welch, eds., *Handbook of Economics of Education, Volume 2*. Amsterdam: Elsevier North-Holland.

Scafidi, Benjamin, David L. Sjoquist, and Todd R. Stinebrickner, 2007. Race, poverty, and teacher mobility, *Economics of Education Review* 26: 145-159.

Statistics Canada, 2008. *Aboriginal Identity (8), Area of Residence (6), Age Groups (12) and Sex (3) for the Population of Canada, Provinces and Territories, 2006 Census - 20% Sample Data.* Statistics Canada Catalogue # 97-558-X2006006, available at http://www.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=97-558-X2006006&lang=eng.

Todd, Petra E. and Kenneth I. Wolpin, 2003. On the specification and estimation of the production function for cognitive achievement: Home, school and racial test score gaps, *The Economic Journal* 113: F3–F33.

Wolfe, Barbara and Robert Haveman, 2001. Accounting for the social and non-market benefits of education, in John Helliwell, ed., *The Contribution of Human and Social Capital to Sustained Economic Growth and Well-being*. Vancouver: University of British Columbia Press.

Tables

Table 1: Per student operating grants to B.C. public school districts.

Category	before March 2002	after March 2002
Base amount	3,042	5,308
Aboriginal supplement	755 – 1,030*	950
ESL supplement	1,230 (Year 1) 1,060 (Years 2-5)	1,100
Special needs supplements:		
Dependent	31,910	30,000
Low incidence/high cost	12,460	15,000
Severe behaviour	6,014	6,000
High incidence/low cost	3,132	0
Gifted	341	0

Source: British Columbia Ministry of Education (2002), page 4.
*amount per student depends on total number of Aboriginal students in the district.

Table 2: Achievement levels and growth, grade 7 students 2002-2004.

Variable	Non-Aboriginal	Aboriginal	Total	Difference
Grade 7 numeracy score				
Mean	0.05	-0.57	0.00	0.62
Standard deviation	0.99	0.87	1.00	
25th percentile	-0.70	-1.20	-0.75	0.51
Median	-0.04	-0.70	-0.09	0.66
75th percentile	0.72	-0.05	0.67	0.77
Grade 7 reading score				
Mean	0.06	-0.58	0.00	0.64
Standard deviation	0.98	1.01	1.00	
25th percentile	-0.62	-1.34	-0.69	0.71
Median	0.08	-0.62	0.03	0.70
75th percentile	0.76	0.15	0.73	0.61
Gain in numeracy score				
Mean	-0.03	-0.08	-0.04	0.05
Standard deviation	0.81	0.76	0.81	
25th percentile	-0.55	-0.57	-0.55	0.02
Median	-0.03	-0.07	-0.04	0.03
75th percentile	0.48	0.42	0.48	0.06
Gain in reading score				
Mean	0.01	-0.08	0.01	0.09
Standard deviation	0.79	0.78	0.79	
25th percentile	-0.48	-0.56	-0.48	0.08
Median	0.03	-0.08	0.02	0.11
75th percentile	0.53	0.43	0.52	0.10

Table 3: Participation in FSA exams, grade 7 students 2002-2004.

Variable	Non-Aboriginal	Aboriginal	Total
% taking grade 7 numeracy exam	90.7	77.2	89.4
% taking grade 7 reading exam	91.6	80.4	90.6
% excused from grade 7 numeracy exam	4.2	10.9	4.8
% excused from grade 7 reading exam	4.2	10.1	4.7
% without numeracy gain data	12.4	28.7	14.0
% without reading gain data	10.9	25.8	12.3

Table 4: Trends in numeracy exam participation, grade 4 and 7 Aboriginal students 1999-2004.

Year	Grade 4 Numeracy		Grade 7 Numeracy	
	% Taking	% Excused	% Taking	% Excused
1999	83.6	7.9	83.6	6.7
2000	82.8	7.8	80.5	9.7
2001	81.7	11.4	77.6	12.8
2002	79.4	9.8	77.9	10.7
2003	78.2	10.8	76.4	11.8
2004	79.0	9.8	74.8	11.1

Table 5: Characteristics of Aboriginal students by participation in grade 7 numeracy exam, 2002-2004.

Variable	Participation in grade 7 exam			
	Took exam	Unexcused absence	Excused absence	Total
Grade 4 numeracy score	-0.5	-0.8	-1.2	-0.5
% excused from grade 4 numeracy exam	3.2	15.2	37.9	8.2
% took grade 4 numeracy exam	91.3	73.7	46.4	84.6
% male	49.0	52.0	61.9	50.8
% with identified disability in grade 7	10.2	28.2	62.4	18.0

Table 6: Bounds on achievement under alternative assumptions about nonparticipants, grade 7 students 2002-2004.

	Bounds on Mean			Bounds on Median		
Variable	**Non-Aboriginal**	**Aboriginal**	**Difference**	**Non-Aboriginal**	**Aboriginal**	**Difference**
Grade 7 Numeracy score						
Worst-case bounds	-	-	-	[-0.17, 0.10]	[-1.00,-0.35]	[0.18,1.10]
Increasing participation	[-0.14, 0.05]	[-0.89,-0.57]	[0.43, 0.94]	[-0.17,-0.04]	[-1.00,-0.70]	[0.53,0.96]
Exogenous participation	0.05	-0.57	0.62	-0.04	-0.70	0.66
Grade 7 Reading score						
Worst-case bounds	-	-	-	[-0.04, 0.20]	[-0.96,-0.28]	[0.24,1.16]
Increasing participation	[-0.12,0.06]	[-0.86,-0.58]	[0.46, 0.92]	[-0.04, 0.08]	[-0.96,-0.62]	[0.58,1.04]
Exogenous participation	0.06	-0.58	0.64	0.08	-0.62	0.70

Table 7: Characteristics of grade 7 students 2002-2004 (percent).

	All Students			Exam Participants Only		
Variable	**Non-Aboriginal**	**Aboriginal**	**Total**	**Non-Aboriginal**	**Aboriginal**	**Total**
# of observations	139,610	14,167	153,777	129,934	11,737	141,671
% of total	90.8	9.2	100.0	91.7	8.3	100.0
% taking numeracy exam	90.7	77.2	89.4	97.4	93.2	97.1
% taking reading exam	91.6	80.4	90.6	98.4	97.0	98.3
Male	51.3	50.8	51.3	50.7	48.9	50.5
Disabled	7.1	18.0	8.1	5.0	11.2	5.5
Physical/sensory disability	1.1	2.2	1.2	0.7	1.1	0.7
Intellectual disability or autism	1.1	3.3	1.3	0.3	0.7	0.3
Severe behavioral disorder	0.8	3.4	1.0	0.5	2.1	0.7
Moderate behavioral disorder	1.1	3.4	1.3	1.0	3.0	1.2
Learning disability	3.0	5.8	3.3	2.5	4.3	2.7
Gifted	2.4	0.7	2.2	2.5	0.8	2.4

"Exam participants" are students that took the grade 7 exam in at least one of the two subjects.

Table 8: Levels regression (dependent variable is grade 7 exam score), all grade 7 students 2002-2004.

Variable	**Numeracy exam**				**Reading exam**			
	(1)	**(2)**	**(3)**	**(4)**	**(1)**	**(2)**	**(3)**	**(4)**
School fixed effects	N	N	Y	Y	N	N	Y	Y
Aboriginal	-0.62***	-0.58***	-0.35***	-0.35***	-0.64***	-0.60***	-0.40***	-0.43***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Male		0.14***	0.15***	0.15***		-0.24***	-0.24***	-0.24***
		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.01)
Learning disability		-0.76***	-0.70***	-0.74***		-0.84***	-0.80***	-0.82***
		(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)
Behavioural disorder		-0.62***	-0.51***	-0.55***		-0.58***	-0.48***	-0.52***
		(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)
Other disability		-0.62***	-0.55***	-0.56***		-0.68***	-0.61***	-0.60***
		(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)
Aboriginal interacted with:								
Male				-0.04**				0.05**
				(0.02)				(0.02)
Learning disability				0.29***				0.17***
				(0.04)				(0.05)
Behavioural disorder				0.20***				0.18***
				(0.04)				(0.05)
Other disability				0.09				-0.06
				(0.07)				(0.07)
Observations	138745	138744	138744	138744	140457	140456	140456	140456
R^2	0.03	0.05	0.19	0.19	0.03	0.08	0.17	0.17

Cluster-robust standard errors in parentheses

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

Year fixed effects in all regressions.

Table 9: Value-added regression (dependent variable is difference between grade 7 and grade 4 exam score), all grade 7 students 2002-2004.

Variable	Numeracy exam				Reading exam			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
School fixed effects	N	N	Y	Y	N	N	Y	Y
Aboriginal	-0.04***	-0.04***	-0.02*	0.00	-0.09***	-0.09***	-0.06***	-0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Male		0.02***	0.02***	0.02***		-0.09***	-0.09***	-0.09***
		(0.00)	(0.00)	(0.00)		(0.01)	(0.00)	(0.01)
Learning disability		-0.01	0.02	0.02		-0.01	0.00	-0.00
		(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)
Behavioural disorder		-0.09***	-0.06***	-0.07***		-0.08***	-0.07***	-0.07***
		(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)
Other disability		-0.07**	-0.04	-0.05*		-0.13***	-0.11***	-0.09***
		(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)
Aboriginal interacted with:								
Male				-0.04**				-0.01
				(0.02)				(0.02)
Learning disability				-0.02				0.04
				(0.06)				(0.06)
Behavioural disorder				0.03				0.03
				(0.04)				(0.05)
Other disability				0.11				-0.10
				(0.08)				(0.08)
Observations	122438	122438	122438	122438	124761	124761	124761	124761
R^2	0.00	0.00	0.12	0.12	0.00	0.00	0.06	0.06

Cluster-robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Year fixed effects in all regressions.

Table 10. Decomposition of Aboriginal/non-Aboriginal test score gaps, all grade 7 students, 2002-2004.

Description	Numeracy exam		Reading exam	
	Grade 7	Growth	Grade 7	Growth
Overall gap	0.619	0.041	0.638	0.089
Fryer-Levitt decomposition:				
Between schools	0.243	0.024	0.216	0.033
Within schools	0.376	0.017	0.423	0.056
Hanushek-Rivkin decomposition:				
Between schools	0.306	0.026	0.287	0.042
Within schools	0.313	0.014	0.351	0.047

Table 11: Peer characteristics of grade 7 students 2002-2004.

Variable	Non-Aboriginal	Aboriginal	Total
% Male peers	51.2	51.7	51.3
	(10.0)	(10.4)	(10.1)
% Aboriginal peers	7.5	26.0	9.2
	(9.5)	(26.3)	(13.2)
% English-language peers	80.0	90.6	80.9
	(25.0)	(17.4)	(24.6)
% ESL peers	5.7	6.4	5.7
	(9.4)	(12.4)	(9.7)
% disabled peers	7.8	11.1	8.1
	(6.5)	(9.2)	(6.9)
% peers with learning/behavioral disability	4.4	5.8	4.5
	(4.5)	(6.2)	(4.7)
% peers with other disability	3.4	5.3	3.6
	(4.0)	(6.3)	(4.3)

Standard deviations in parentheses

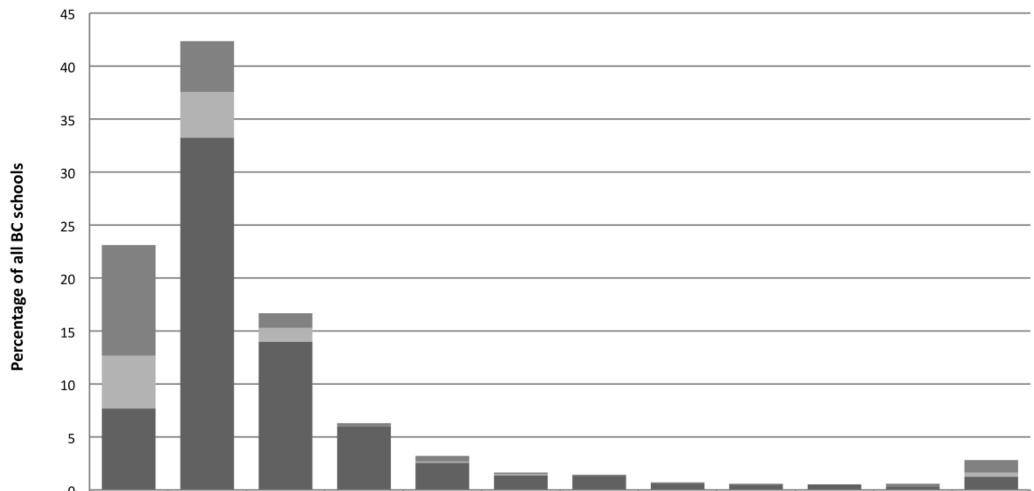
Table 12: Value-added regression for effect of peer background characteristics (dependent variable is difference between grade 7 and grade 4 exam score), grade 7 Aboriginal students 2002-2004.

Variable	Numeracy exam			Reading exam		
	(1)	(2)	(3)	(1)	(2)	(3)
Individual variables:						
Male	-0.03	-0.03	-0.02	-0.11***	-0.10***	-0.10***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Disabled		-0.01			-0.04	
		(0.03)			(0.03)	
Learning disability			-0.01			0.05
			(0.06)			(0.06)
Behavioural disorder			-0.03			-0.06
			(0.04)			(0.04)
Other disability			0.09			-0.21***
			(0.08)			(0.08)
Peer variables:						
% male peers	-0.19	-0.18	-0.18	-0.12	-0.13	-0.13
	(0.11)	(0.12)	(0.12)	(0.10)	(0.10)	(0.10)
% Aboriginal peers	0.26	0.26	0.24	0.08	0.07	0.07
	(0.16)	(0.16)	(0.16)	(0.14)	(0.14)	(0.14)
% disabled		-0.03			0.12	
		(0.18)			(0.19)	
% learning/behavioral disability			-0.23			0.09
			(0.23)			(0.24)
% other disability			0.22			0.14
			(0.26)			(0.26)
Total peer effect	0.047	0.046	0.045	0.015	0.016	0.017
	(0.029)	(0.030)	(0.030)	(0.026)	(0.026)	(0.026)
Observations	9611	9611	9611	10006	10006	10006
R^2	0.21	0.21	0.21	0.17	0.17	0.17

Cluster-robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
School and year fixed effects in all regressions.
"Total peer effect" is the predicted change in outcome implied by changing from the average peer group of non-Aboriginal students to the average peer group of Aboriginal students. See text for details.

Figures

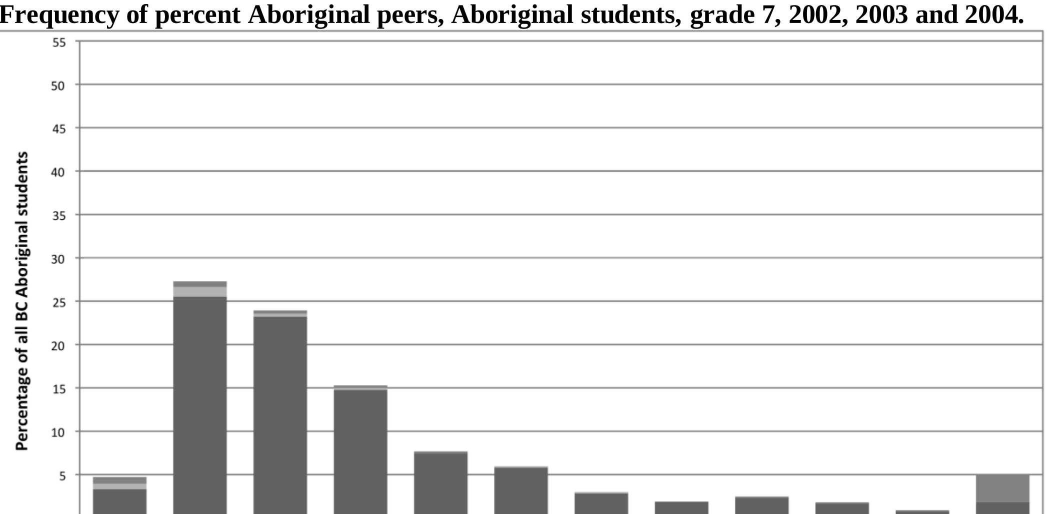
Figure 1. Frequency of school percent Aboriginal, grade 7, 2002, 2003 and 2004



	0	0<x<10	10≤x<20	20≤x<30	30≤x<40	40≤x<50	50≤x<60	60≤x<70	70≤x<80	80≤x<90	90≤x<100	100
PRIVATE	10.44	4.79	1.38	0.33	0.53	0.13	0.07	0.07	0.07	0.00	0.26	1.18
FRENCH IMMERSION	4.99	4.33	1.31	0.00	0.07	0.07	0.00	0.00	0.00	0.00	0.00	0.39
REGULAR PUBLIC	7.68	33.22	13.99	5.98	2.63	1.44	1.38	0.66	0.53	0.53	0.33	1.25

Percentage of Aboriginal students

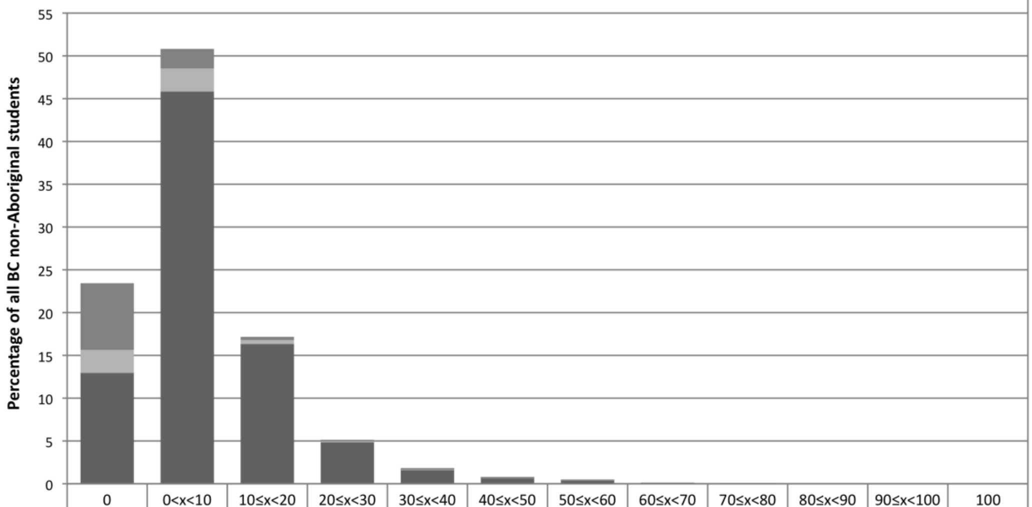
Figure 2. Frequency of percent Aboriginal peers, Aboriginal students, grade 7, 2002, 2003 and 2004.



	0	0<x<10	10≤x<20	20≤x<30	30≤x<40	40≤x<50	50≤x<60	60≤x<70	70≤x<80	80≤x<90	90≤x<100	100
PRIVATE	0.78	0.64	0.37	0.29	0.24	0.00	0.00	0.00	0.03	0.11	0.06	3.22
FRENCH IMMERSION	0.62	1.11	0.34	0.22	0.00	0.02	0.04	0.00	0.00	0.00	0.00	0.00
REGULAR PUBLIC	3.32	25.54	23.23	14.78	7.44	5.95	2.95	1.88	2.44	1.69	0.84	1.85

Percentage of Aboriginal peers

Figure 3. Frequency of percent Aboriginal peers, non-Aboriginal students, grade 7, 2002, 2003 and 2004.



	0	0<x<10	10≤x<20	20≤x<30	30≤x<40	40≤x<50	50≤x<60	60≤x<70	70≤x<80	80≤x<90	90≤x<100	100
PRIVATE	7.80	2.28	0.37	0.13	0.08	0.01	0.01	0.00	0.00	0.00	0.00	0.01
FRENCH IMMERSION	2.68	2.70	0.47	0.07	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
REGULAR PUBLIC	12.96	45.84	16.33	4.93	1.71	0.78	0.49	0.12	0.10	0.04	0.02	0.02

Percentage of Aboriginal peers

Not-for-publication appendix

This appendix provides additional results and robustness checks in support of the main paper.

A1 Robustness of results to handling of incomplete exam data

Some students participate in some exams but not others. In the main text, we have chosen to include students in our data whenever possible. For example, Table 2 reports the mean grade 7 numeracy score for all students who took that exam, whether or not they took the other exams. An alternative is to report results for a common sample of students with complete data.

Table A1 reproduces the summary statistics reported in Table 2, but with the sample restricted to only those students that took both exams in both grades. The results are similar to those in Table 2. As one might expect, scores are slightly higher among students who take all exams than among students that miss at least one exam. In addition, the gap in grade 7 reading scores is slightly smaller.

Table A2 reproduces the levels regression of Table 8 using a sample restricted to those students who took the exam in both grades, and thus appeared in the value-added regressions of Table 9. The results are similar to those reported in Table 8.

A2 Additional bounds

Table A3 reports bounds on test score gains, calculated in a manner similar to the bounds on grade 7 test scores reported in Table 6. Table A3 also reports bounds on test scores and test score gains using the method developed by Lee (2009).

The worst case and exogenous participation bounds are calculated as in Table 6. The increasing participation bounds are calculated under the assumption that the test score gain of students with missing gain scores is bounded below by -2.0 standard deviations, and that the probability of participation is increasing in the achievement gain. As Table A3 shows, the bounds on test score gains tend to be substantially less informative than the corresponding bounds on test score levels.

Table A3 also reports bounds on average test scores and the test score gap using the method developed by Lee (2009). In contrast to the conventional approach of making assumptions about the unobserved outcomes of nonparticipants, Lee argues for comparing the outcomes of participants in the low-participation group to the outcomes of participants in a trimmed sample of the high-participation group. The proportion of cases to be trimmed is $P = \frac{participation_H - participation_L}{participation_H}$, where $participation_H$ is the participation rate in the high-participation group, and $participation_L$ is the participation rate in the low-participation group. The lower bound for mean/median in the high participation group is estimated by trimming the top P of scores, while the upper bound is estimated by trimming the bottom P. For example we compare the average numeracy score of the 77% of Aboriginal students who took the numeracy exam to the range of average numeracy scores we would get after trimming the non-Aboriginal sample to have a participation rate of 77%.

We can compare the Lee bounds in Table A3 to those reported in Table 6 under the assumption that participation rates are increasing in achievement. The Lee bounds on the mean and median non-Aboriginal test score are substantially wider than the corresponding increasing participation bounds. However, the Lee bounds on the test score gap are quite similar to the increasing participation bounds because the sample for the low-participation group (i.e. Aboriginal students) is not truncated at all.

While this calculation is straightforward, its interpretation is somewhat more complex. Lee's method was developed in a treatment effects framework, in which the two groups are "treatment" and "control." In that setting, Lee's calculation can be interpreted as bounding the average treatment effect among individuals whose participation is unaffected by the treatment. This interpretation does not apply to our setting, as the relevant counterfactual – the exam participation that would be observed for a given non-Aboriginal student if that student were Aboriginal – makes little sense. The Lee bounds could be interpreted as the range of observed achievement we might see if non-Aboriginal students were to have the same participation rate as Aboriginal students. However, they do not necessarily bound any specific population parameter of interest.

Appendix Tables

Table A1: Achievement levels and growth, grade 7 students with complete outcome data 2002-2004.

Variable	Non-Aboriginal	Aboriginal	Total	Difference
Grade 7 numeracy score				
Mean	0.06	-0.50	0.02	0.66
Standard deviation	0.98	0.87	0.98	
25^{th} percentile	-0.67	-1.14	-0.72	0.47
Median	-0.03	-0.63	-0.08	0.60
75^{th} percentile	0.72	0.01	0.67	0.71
Grade 7 reading score				
Mean	0.11	-0.47	0.07	0.58
Standard deviation	0.96	0.99	0.97	
25^{th} percentile	-0.55	-1.19	-0.61	0.64
Median	0.13	-0.49	0.09	0.62
75^{th} percentile	0.80	0.25	0.76	0.55
Gain in numeracy score				
Mean	-0.03	-0.08	-0.04	0.05
Standard deviation	0.81	0.76	0.81	
25^{th} percentile	-0.55	-0.57	-0.55	0.02
Median	-0.03	-0.07	-0.04	0.04
75^{th} percentile	0.48	0.42	0.48	0.06
Gain in reading score				
Mean	0.01	-0.07	0.01	0.08
Standard deviation	0.79	0.77	0.79	
25^{th} percentile	-0.47	-0.54	-0.48	0.07
Median	0.03	-0.07	0.02	0.10
75^{th} percentile	0.53	0.43	0.52	0.10

Table A2: Levels regression (dependent variable is grade 7 exam score), grade 7 students with value-added data 2002-2004.

Variable	Numeracy exam (1)	(2)	(3)	(4)	Reading exam (1)	(2)	(3)	(4)
School fixed effects	N	N	Y	Y	N	N	Y	Y
Aboriginal	-0.58***	-0.55***	-0.33***	-0.33***	-0.61***	-0.59***	-0.39***	-0.42***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)
Male		0.14***	0.15***	0.15***		-0.25***	-0.24***	-0.24***
		(0.01)	(0.01)	(0.01)		(0.01)	(0.01)	(0.01)
Learning disability		-0.73***	-0.67***	-0.70***		-0.81***	-0.78***	-0.79***
		(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)
Behavioural disorder		-0.59***	-0.48***	-0.52***		-0.57***	-0.47***	-0.51***
		(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)
Other disability		-0.52***	-0.46***	-0.47***		-0.58***	-0.51***	-0.51***
		(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)
Aboriginal interacted with:								
Male				-0.04**				0.05**
				(0.02)				(0.02)
Learning disability				0.20***				0.16***
				(0.04)				(0.05)
Behavioural disorder				0.22***				0.18***
				(0.04)				(0.05)
Other disability				0.13				-0.07
				(0.09)				(0.09)
Observations	122438	122438	122438	122438	124761	124761	124761	124761
R^2	0.02	0.05	0.19	0.19	0.03	0.07	0.17	0.17

Cluster-robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Year fixed effects in all regressions.

Table A3: Additional bounds on achievement, grade 7 students 2002-2004.

Variable	Bounds on Mean: Non-Aboriginal	Aboriginal	Difference	Bounds on Median: Non-Aboriginal	Aboriginal	Difference
Grade 7 Numeracy score						
Lee bounds	[-0.24, 0.30]	-0.57	[0.33, 0.86]	[-0.23, 0.16]	-0.70	[0.46,0.86]
Numeracy score gain						
Worst-case bounds	-	-	-	[-0.17, 0.10]	[-0.46, 0.31]	[-0.49, 0.56]
Increasing participation	[-0.28,-0.03]	[-0.63,-0.08]	[-0.20, 0.59]	[-0.17,-0.03]	[-0.46,-0.07]	[-0.11, 0.42]
Exogenous participation	-0.03	-0.08	0.04	-0.03	-0.07	0.03
Lee bounds	[-0.20, 0.23]	-0.08	[-0.20, 0.28]	[-0.19, 0.12]	-0.07	[-0.13, 0.19]
Grade 7 Reading score						
Lee bounds	[-0.16, 0.29]	-0.58	[0.42, 0.87]	[-0.08, 0.24]	-0.62	[0.54, 0.86]
Reading score gain						
Worst-case bounds	-	-	-			
Increasing participation	[-0.21, 0.01]	[-0.57,-0.08]	[-0.13, 0.59]	[-0.09, 0.03]	[-0.40,-0.08]	[-0.01, 0.43]
Exogenous participation	0.01	-0.08	0.09	0.03	-0.08	0.11
Lee bounds	[-0.20, 0.23]	-0.08	[-0.12, 0.31]	[-0.11, 0.16]	-0.08	[-0.03, 0.24]

Bounds on test score gains are calculated for students attending B.C. schools for both grade 4 and grade 7.

Appendix References

Lee, David S., 2009. Training, wages, and sample selection: Estimating sharp bounds on treatment effects, *Review of Economic Studies*, forthcoming.