DISCUSSION PAPER SERIES

IZA DP No. 4824

Differences in the Distribution of High School Achievement: The Role of Class Size and Time-in-Term

Miles Corak
Darren Lauzon

March 2010

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Differences in the Distribution of High School Achievement: The Role of Class Size and Time-in-Term

Miles Corak
University of Ottawa,
Statistics Canada and IZA

Darren Lauzon
Statistics Canada

Discussion Paper No. 4824
March 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 4824
March 2010

ABSTRACT

Differences in the Distribution of High School Achievement: The Role of Class Size and Time-in-Term*

This paper adopts the technique of DiNardo, Fortin and Lemieux (1996) to decompose differences in the distribution of PISA test scores in Canada, and assesses the relative contribution of differences in the distribution of "class size" and time-in-term, other school factors and student background factors. Class size and time-in-term are both important school choice variables and we examine how provincial achievement differences would change if the Alberta distribution of class size and time-in-term prevailed in the other provinces. Results differ by province, and for provinces where mean achievement gaps would be lower, not all students would benefit.

JEL Classification: I22

Keywords: educational economics, human capital, input-output analysis

Corresponding author:

Miles Corak
Graduate School of Public and International Affairs
University of Ottawa
Desmarais Building, Room 11156
55 Laurier Avenue East
Ottawa, Ontario K1N 6N5
Canada
E-mail: corak@iza.org

* Earlier versions of this paper were presented at Symposium 2002: Using Survey Data for Social and Economic Research (Ottawa) and the Canadian Employment Research Forum session "Education Schooling and the Labour Market," Canadian Economics Association meetings May 29-30 2003, Carleton University, Ottawa. The authors thank participants at these events for useful comments and suggestions. All errors and omissions remain our responsibility. The opinions expressed are those of the authors and do not necessarily reflect the opinions of Statistics Canada.

I. Introduction

Variation in school outcomes across jurisdictions raises important questions about equal access to good education and the effectiveness of policy alternatives. Provincial variation has been a consistent feature of recent assessments of high school achievement in Canada over the 1990s.[1] Though not directly comparable, the general pattern in these assessments is one in which the central and western provinces do better, according to summary measures like mean test scores, than the eastern provinces. Québec and Alberta consistently perform well in tests of mathematics and science. Results from the 2000 Programme for International Student Assessment (PISA) confirm this pattern.

It is natural to ask to what extent provincial variation in school inputs and the composition of student populations contributes to these differences. The reviews by Hanushek (1986, 1996) show that in many observational studies, the variation in test scores explained by school inputs such as per pupil funding, class size and teacher qualifications is small relative to family background (usually family income) and other "home environment" variables. School effects tend to be small, statistically insignificant and often the wrong sign. The implication of this is that variation in student backgrounds across provinces would explain more of the observed provincial gaps. Hanushek and Taylor (1990) find this in a study of variation between U.S. states in educational outcomes.

Card and Krueger (1992) point out that this is in contrast to the literature that shows school inputs are important for explaining labour market outcomes such as earnings and employment and that there are too many positive findings in the literature to be due merely to chance. They cite meta-analyses such as those of Hedges, Laine and Greenwald (1994) that come to different conclusions reviewing the same studies as Hanushek. Loeb and Bound (1996) find evidence that cohorts and data aggregation explain the divergence in the two literatures; " studies finding positive effects of school inputs typically use aggregate data on cohorts educated before 1960 while studies finding no effects tend to use micro-level data on more recent cohorts". Recent developments in the program evaluation and treatment effects literature underscore the data and estimation requirements for identifying genuine treatment effects for

[1] See Corak and Lauzon, 2002.

policy alternatives which are difficult if not impossible to implement with most observational studies.[2] For example, where improvements in achievement from class-size reduction have been found they have tended to come from experimental studies or from studies with genuinely exogenous variation in class-size (Lazear, 1999).

Our paper is motivated by two observations. First, school inputs are choice variables and so observed levels reflect in part the features of the student population. This is not often accounted for in studies of student achievement and simply adding student and school input variables into a linear regression model can be problematic. Lazear (1999) shows that puzzling findings in the literature on class size can be explained by considering a simple public-goods model of classroom learning and its implications for optimal class size. Second, the vast majority of papers employ variations of a linear parametric regression model of an education production function. Specification issues involving such models have been well documented (Hanushek, 1979, 1986). Recent panel data in the U.S. have allowed improved estimation, particularly by allowing one to estimate value-added models so that "innate ability" is better controlled for, but omitted variables and measurement error remain important. These are especially relevant in the context of random effects and higher-order error components models, like the hierarchical linear model (HLM), where strong assumptions about the correlation between omitted variables and included regressors are required to obtain desirable statistical properties.[3]

Despite these concerns, surprisingly few studies have addressed the use of the linear parametric regression framework and its emphasis on average outcomes. School systems do not uniformly impact students. If the role of schools is to bring all students to a minimum standard of achievement, regardless of initial cognitive endowments, the expected impact of schooling inputs would be greater for the least skilled (or most at-risk) students. For example, where smaller sizes have had a positive effect, it has tended to be greatest for disadvantaged students (Lazear, 1999). Alternatively, school resources may be more productively used on students with greater learning potential. Vulnerable students may fall through the cracks and so changes in the school inputs would have greater effect for higher achievers. Some papers have examined the distribution of school outcomes (Levin, 2000, Beddard, Brown and Helland, 1999, Eide

[2] See Heckman and Vytlacil (2005) for a detailed summary and discussion of the various approaches and a reformulation of them in terms of the marginal treatment effect.

and Showalter, 1998) but continue the use of parametric forms (quantile regression in the case of Eide and Showalter and Levin and ordered probit in the case of Beddard et. al.).

Lastly, many studies emphasize the marginal effects of school inputs on an achievement variable. Notwithstanding the difficulties in obtaining valid treatment effects parameters, though the effect of a small change in class size might be small, large variation between policy jurisdictions in the distribution of class sizes could contribute a large amount to observed differences in achievement. They could even contribute more than the variation in family background characteristics if these do not differ substantially between regions even if family background characteristics have larger marginal effects. The same is true of variations in achievement over time. Cook and Evans (2000) find a small contribution for both student background and school factors in explaining the black – white achievement convergence in the U.S. In such cases, decomposition of differences provides another source of evidence on the relative importance of school and student inputs.

In this paper we use the semi-parametric approach developed by DiNardo, Fortin and Lemieux (DFL, 1996) to decompose differences in provincial achievement distributions into components attributable to student background and school factors. Their approach has two advantages for our purposes. First it is a direct method of decomposition that bypasses the need to identify valid treatment effects in data that may not support their identification. Second, the approach is simpler than the methods discussed in Doiron and Riddell (1994) and Fairlee (1999) for decomposition designed for models that are non-linear in parameters. This paper extends previous work (Corak and Lauzon, 2002) which explored the relative role of school inputs and student background factors in explaining provincial variation in achievement distributions using the 1999 PISA data for Canada. We focus here on the role of "class-size" and time-in-term, the latter referring to the total amount of instructional time and how it is distributed through the year. These are important control variables for school administrators and reflect choices of teaching input. Much has been written on class size and public debate on the merits of class size reduction continues. Reductions cost money but (it is argued) the benefit is higher student performance. Reallocating time-in-term is also becoming a serious policy option and is being discussed in the popular press. Recently, school districts in

[3] For examples of the HLM in education research see Raudenbush and Bryk (1986), Willms and Raudenbush (1989) and Raudenbush and Willms (1995) .

the U.S. and Canada (the Grand Forks District in BC) have adopted four day school weeks and report significant cost savings and increased student performance.

We examine differences between Alberta, the highest performing province in the PISA assessments, and the Atlantic provinces in reading, mathematics and science achievement distributions. The Atlantic provinces are the only provinces whose mean scores are statistically significantly lower than Alberta's in all three subject domains. Two provinces, Nova Scotia and New Brunswick also have dual language school systems. We therefore do a separate analysis for these provinces for the English language sector (the majority sector in both) and use this to infer the contribution of within province differences between the language sectors to the gap with Alberta. We get quite different results depending on which provincial difference we analyze, for what part of the distribution and for which subject domain. In some cases differences in class size and time-in-term do not explain differences in mean or median performance because of offsetting changes in the upper and lower parts of the achievement distribution. In most cases it is clear that differences in class size and time-in-term do account for differences in mean or median performance but that this can mask the fact that these differences might actually reduce achievement differences between the provinces in particular parts of the distribution. For example, we find that differences in the distribution of class size and time-in-term between Alberta and New Brunswick explain a large part of the difference in average reading performance between these two provinces but these differences actually reduce differences in the proportion of students performing below the lowest proficiency standard in reading. This observation is an important one. Policy makers influenced by an analysis of average differences might seek policies that could disadvantage the most vulnerable students.

In our data, Alberta has the largest class-size and more instructional time per year distributed in more classes of shorter duration per week. As we discuss in section II and III, we are not here trying to obtain the effects on achievement of reducing class size or time-in-term. Indeed observational studies tend to show that larger classes have better achievement while experimental studies tend to show the opposite (see section II). Here, we are interested strictly in how the distribution of test sores would look in other provinces if they had the same distribution of class size and time-in-term as Alberta.

The paper is organized as follows. Section II discusses some recent literature on class size reduction and time-in-term. Section III details our use of the DFL decomposition approach. Section IV describes the Canadian PISA data and the factors we consider. Results are in Section V and VI, Section VII concludes.

II. Class size and time-in-term

The debate in the large literature on class sizes centres on the extent to which smaller classes improve student achievement, and whether this benefit (if it exists) is cost effective. Hanushek (1998) argues that evidence about the achievement improvements from smaller class sizes is "meager and unconvincing". Referring to the STAR experiment in Tennessee, he further argues that "widely cited experimental evidence actually offers little support for reductions in class size." Examining the data from the same experiment, Krueger (1997) concludes the opposite.[4] Ehrenberg et. al. (2001) find that in other studies, "quasi experimental" findings tend to support the Tennessee results. Debate is still open. Even studies that carefully identify truly exogenous variation in class-size can find different results; Hoxby (2001) in the negative and Angrist and Lavy (1997) in the positive.

Lazear (1999) suggests that many of these "puzzling" findings in the literature can be explained by a simple model of education production that treats classroom learning as a public good. Consumption of the public good is disrupted if the teacher must focus attention on an individual student. This can occur not only if a student is disruptive, but if a student asks a question to which everyone else in the class knows the answer. If p represents the probability that any one student is not disrupting the class at a given time, then the public good is being consumed with probability p^n in a class of size n. This is the key parameter of the Lazear model and the comparative statics of this simple model reveal that optimal class size varies directly with this (smaller p implies a higher optimal level of teachers which in turn implies smaller class sizes). Inconclusive test score results reported in the literature can follow from two sources. First, the magnitude of any improvement in overall class learning from a reduction in class size can be quite small depending on the values of p, the costs of teacher time, the productivity of a moment of teacher time, the returns to a moment of teacher time and the current class size. Second, because p is negatively related to the optimal choice of teaching input, smaller classes have students with lower p in them (i.e., have more "disruptive" students). Lazear shows that the positive effect of reducing class size is insufficient to overcome this

[4] Hanushek did not have access to the data but confined his remarks to published reports about the STAR experiment.

deficiency. For this reason, positive class size effects in cross-sectional studies are not so surprising. More importantly, these inconclusive results do not mean there is no *potentially large* class size effect. If any group of students with a fixed p were placed in a larger class, educational output would fall. This is why experiments that leave p constant find expected results (as examples, he cites Krueger, 1997 and Angrist and Lavy, 1997).[5] Hoxby (2001), however, did not find any significant advantage to class size reduction in her study of a natural experiment in Connecticut.

Lazear's model also introduces the important idea that class room learning is a public good. This conceptually connects the notion of class size with time-in-term. Class size in a public goods setting refers to the number of "consumers" of the good. Time-in-term refers to the total amount of the good that is available to consume.[6] Time-in-term is typically not studied because of little variation in most available data sets. Schools have begun considering both the total amount of instructional time and how it is allocated. In Arizona, several school districts have switched to a four day school week, as has The Grand Forks district in BC. In the context of the Lazear model, both class size and time-in-term are directly related to the choice of teacher inputs. It makes sense to treat these inputs as closely related. In our analysis, we treat time-in-term and class-size as all describing the teaching input in the school. In the spirit of the Lazear model, these are important policy levers that schools can use to manage costs while delivering education. This allows for two schools with the same class size to have a different effect if they allocate teaching time differently. Thus we do not compare the relative contribution of the two to achievement differences but consider them in total.

Finally it should be noted that many studies use the student-teacher ratio rather than an actual count of students in a particular class and this is often criticized as not reflecting the experience of individual students whose performance is the focus of study. Individual class sizes vary within schools for a number of reasons. One important reason is that secondary school students typically take different subjects and these are sometimes taught in multiple class rooms by different teachers. Thus, class size as reported by

[5] Hoxby (2001) suggests that Hawthorne effects and other "reactive behaviour" could explain why her results differ from those of policy experiments like STAR; "school administrators "make good use of smaller class sizes because full enactment of the policy depends on a successful evaluation". In her data, school staff were unaware of the natural experiment.

[6] Alternatively, if longer academic years are devoted to a larger array of topics, rather than more detailed attention paid to the same topics taught in shorter academic years, there is a greater spectrum of related public goods to consume.

students in a particular classroom is a result not only of optimal teaching inputs determined by school administrators but also the selection of particular classes by students. The relevant class size to relate to a particular achievement outcome in this case is not always clear and selection by students should be accounted for. Student-teacher ratios are a useful measure of the overall amount of teaching resources per student in the school. The kind of variable that is most appropriate depends on the intended analysis. This study considers variations in class size to be a reflection of variations in the optimal level of an input variable so the student-teacher ratio is the desired measure. It turns out that there is more interprovincial variation in this variable than in student-reported class size.

III. The DFL decomposition

This section describes the DFL approach in the context of achievement differences between provinces. Let (y, p, z) be a jointly distributed random vector of test scores, provinces, and school and family background covariates respectively. The goal is to decompose the difference in marginal test score densities $f_1(y) - f_0(y)$ into parts attributable to differences between the two provinces in the distribution of different components of z. Therefore, we require counterfactual density functions that depict the distribution of test scores in a given province (province 1) if the school or family background characteristics were distributed as they are in a baseline province (province 0) and students are otherwise educated as they would be in province 1. This is a generalization of the familiar Oaxaca decomposition (Oaxca, 1973) to differences in distributions. The central insight of DFL is that these counterfactual densities are obtained simply by re-weighting the actual density.

The starting point is to first note that marginal test score densities we want to difference can be expressed as the product of the probability of observing a test score conditional on a given value of the covariates in province i times the probability of observing the covariate values in province i:

$$\begin{aligned} f_i(y) &= \int_{z\in\Omega} dF(y, z \mid p_y = p_z = i) \\ &= \int_{z\in\Omega} f(y \mid z, p_y = i)\, dF(z \mid p_z = i) \\ &= f(y; p_y = p_z = i) \\ i &= 1,0 \end{aligned} \tag{1}$$

where Ω is the covariate support. The last line simply makes it explicit that we consider several provinces at a time, the province whose test score distribution we wish to obtain, and the province whose covariate distribution we use to obtain it.

The counterfactual densities can be expressed in terms of the data from province 1 using a weight adjustment defined as follows:

$$\begin{aligned} f_i(y) &= \int_{z\in\Omega} f(y \mid z, p_y = 1)\, dF(z \mid p_z = 0) \\ &= \int_{z\in\Omega} f(y \mid z, p_y = 1)\,\psi(z)\, dF(z \mid p_z = 1) \\ \psi(z) &\equiv \frac{dF(z \mid p_z = 0)}{dF(z \mid p_z = 1)} \end{aligned} \tag{2}$$

With many covariates, it is difficult or impossible to identify all the probabilities necessary to compute the weighting function $\psi(z)$. However using Bayes rule it can be re-expressed in terms of the probabilities of a single variable (the province) conditional on the covariates.

$$\psi(z) = \left(\frac{pr(p_x = 0 \mid z)}{pr(p_x = 1 \mid z)}\right)\cdot\left(\frac{pr(p_x = 1)}{pr(p_x = 0)}\right) \tag{3}$$

The first ratio is just the odds of being in province 0 conditional on the covariates and the second is the (marginal) odds of being in province 1. Both are directly estimable using the logit model.

The procedure can best be understood with reference to a simple example. Tables A, B and C show hypothetical data for two provinces “E” and “W”, a dichotomous test score variable with values “H” and “L” and a dichotomous covariate *z* with values 0 and 1. The tables are constructed to show that the test score difference between “E” and “W” is due largely to the distribution of the covariate *z* within each province. Persons with *z* = 1 tend to have a low score and there are more persons with *z* = 1 in province “E”. The first equality in (1) simply states that we obtain the row marginals within province by summing the cell proportions across columns. The second equality states that we can also obtain the row marginals by multiplying the column proportions of *z* by the column marginals and summing over the columns. For example, prob(y = H | z = 0, $p_z = p_y$ = E) = (0.5 x 0.3158) + (0.0769 x 0.6842) = 0.2105.

Table A: A Hypothetical Data Set

Test Score	Province	z	Freq	Weighting Function	Adjusted Frequency
H	E	0	30	2.25	67.55
H	E	1	10	0.42	4.22
H	W	0	150	.	.
H	W	1	25	.	.
L	E	0	30	2.25	67.55
L	E	1	120	0.42	50.67
L	W	0	10	.	.
L	W	1	40	.	.

Table B: Conditional Distributions of the Hypothetical Data

Frequency Percent Row Pct Col Pct	Province = E			Frequency Percent Row Pct Col Pct	Province = W		
	0	1	Total		0	1	Total
H	30 0.1579 0.75 0.5	10 0.0526 0.25 0.0769	40 0.2105	H	150 0.6667 0.8571 0.9375	25 0.1111 0.1429 0.3846	175 0.7778
L	30 0.1579 0.20 0.5	120 0.6316 0.80 0.9231	150 0.7895	L	10 0.0444 0.2 0.0625	40 0.1778 0.8 0.6154	50 0.2222
Total	60 0.3158	130 0.6842	190 1	Total	160 0.7111	65 0.2889	225 1

It is not necessary to apply the DFL weight adjustments to these simple tables to obtain the counterfactuals because the number of cells is so small that all the probabilities are easily identified. So we can simply compute the probabilities directly from the first equation in (2). For example, to compute prob(y = H | z = 0, p_z = W, p_y = E) we simply compute (0.5 x 0.7111) + (0.0769 x 0.2889) = 0.3778 This shows a sizeable increase in the proportion performing at the high level in province E if in E, *z* had the distribution it does in W, as we would expect. To apply the DFL method, we can use (2) to directly compute the adjustments, which is in this example are prob(z = 0 | p_z = W)/prob(z = 0|p_z = E) = 0.7111/0.3158 and prob(z = 0 | p_z = W)/prob(z = 0|p_z = E) = 0.2889/0.6842 and insert them in the preceding calculation, canceling out 0.3158 and 0.6842. However, we can also reweight the data for province E by the amount z x (0.2889/0.6842) + (1 – z) x (0.7111/0.3158) and multiply this by the frequency in each cell. Thus, for the frequency in province E with test score = H and z = 0, we have 30 x ((0.7111/0.3158)) = 67.552 and for the frequency in province E with test score = H and z = 1 we have 10 x (0.2889/0.6842) = 4.224. Table A shows the adjustment we make to the data and Table C shows the adjusted distributions for province E. We see that the marginal probability of being high scoring in province E is equal to 0.3778 as computed directly above.

Table C: The counterfactual Distribution for Province E

Frequency Percent Row Pct Col Pct	0	1	Total
H	67.552 0.3555 0.9412 0.5	4.2224 0.0222 0.0588 0.0769	71.775 0.3778
L	67.552 0.3555 0.5714 0.5	50.669 0.2667 0.4286 0.9231	118.22 0.6222
Total	135.104 0.7111	54.8918 0.2889	189.996 1

In DFL and in this study, the marginal densities are estimated by means of the kernel density estimator:

$$\widehat{f}_1(y) = \sum_{i \in p_y = 1} \frac{\theta_i}{h} K\left(\frac{y_0 - y_i}{h}\right) \equiv \kappa(K, h, \theta_i) \qquad (1).$$

The kernel density estimator has been discussed in several papers (DiNardo, Fortin and Lemieux, 1996; Blundell and Duncan, 1998; Yatchew, 1998; DiNardo and Tobias (2001)). Here θ_i is the sample weight, normalized to sum to one. The function *K* is the kernel and gives decreasing weight to points of greater distance from y_0. The kernel estimator is a generalization of the familiar histogram which can be obtained from (2) with a suitably chosen kernel. Generally, estimates are robust to choices of *K* but not to different choices of *h*.[7] The tradeoff is one of variance versus bias. If *h* is too large, the density will be over-smoothed relative to the true density and if *h* is too small, the true shape of the density will be estimated imprecisely. The choice of *h* remains an open subject of research. DFL use the "plug-in" method of

[7] Restricting *K* to a certain class of functions.

Sheather and Jones (1991) as this has been shown to be a better selection in cases of complex, multi-modal densities (Park and Turloch, 1992). Since the underlying plausible value estimates used in this paper are drawn from symmetric probability distributions, this is less of a concern with this data. In this study, we use the "rule-of-thumb" estimator suggested by Silverman (1986), $h = 0.09(\min\{\hat{\sigma}, IQR / 1.34\}) n^{-1/5}$, where $\hat{\sigma}$ is the sample standard deviation, *IQR* is the inter-quartile range (the difference between the 75th and 25th percentiles) and *n* is the sample size.

The counterfactual density is obtained by first estimating the ratios in (3) using the logit model and then using them to compute a weight adjustment for each individual in province 1. The counter factual density is obtained using the kernel density estimator with $\theta_i^{'} = \theta_i \psi(z)$ as the new weight. In the appendix we show the computation of the weighting functions and the definition of the weights to use for the density estimation when the covariate vector *z* is partitioned into the components considered in this study.

We emphasize here that the DFL approach is a direct method of decomposing differences in density functions, bypassing the usual approach of estimating a model of conditional means or probabilities and using this to obtain the counterfactual simulations. It is not a means of estimating a conditional density function. The logit is not used to identify the marginal effects of school or student variables on test scores. It is used to estimate the predicted probability that a student with a given set of characteristics is in a given province. Omitted variables are not as critical here since omitting certain variables from Z simply means we "integrating out" the unobserved covariates to obtain $prob(p_z = i \mid Z = z)$ just as we could obtain the marginal probability $prob(p_z = i)$ by integrating out the remaining Z. Donald, Green and Paarsch (2000) propose a method of estimating conditional distribution functions with marginal effects and as indicated in the introduction, ordered probit and quantile regression have been employed with test score data. In our view, the DFL approach has two important advantages over these alternatives. First, as both the Hanushek reviews of the issues involved with estimating the education production function and the recent literature on treatment effects have shown, exceptional data that contains not only the relevant variables but genuinely exogenous variation in the policy variables and covariates is critical to obtain valid estimates of the true treatment effects of school policy alternatives. The DFL approach provides a means of assessing the contribution of differences in observables to differences in test scores in observational data without entering into the estimation of treatment effects that the data cannot support. The second advantage we see is that the direct decomposition approach avoids the complexities of decomposing differences in probabilities with models that are nonlinear in the parameters such as those discussed by Doiron and Riddell (1994) and by Fairlee (1999). The obvious disadvantage is that we cannot obtain estimates of marginal effects with DFL. Again, given the challenges of finding valid treatment effects in observational data, we feel the advantages outweigh this disadvantage.

In the decompositions that follow we assess differences in school characteristics before controlling for differences in student background. In particular we consider first differences in class-size and time-in-term, then other school factors then student back ground characteristics. We do this because the policy experiment we have in mind is what would happen if the (apparently) successful school characteristics of Alberta were adopted by other provinces *given their specific populations*. One of our primary hypotheses is that school systems organize themselves in a way that is optimal given the student population they must serve. We thus consider first the key choice variables under consideration, holding constant other school factors and student characteristics, then allow the others to change in turn. We recognize though, that the effects we observe are sensitive to the order of decomposition so we report results for the reverse order decomposition (as do DFL, 1996) in section V.6.

IV. The data

We use the Canadian results from the Programme for International Student Assessment (PISA) conducted by member OECD countries in April and May of 2000. The PISA is based on a two-stage design with about 1200 schools sampled at a first stage then a random sample of 15 year old students within the schools taken in the second stage.[8] Students were administered a two-hour written test to assess their knowledge of reading, mathematics and science. The primary subject domain of the 2000 PISA was reading meaning that about two-thirds of the test items were reading related. Surveys were administered to students who participated in the test as well as principals of their schools. In Canada, the PISA was integrated with the Longitudinal Youth in Transition Survey (YITS), so participating students also

[8] In the final data set there were 1117 schools for the reading and science assessments and 1116 schools for the mathematics assessment. There were 29,687 students for the reading assessment and 16,489 students for the mathematics and science assessments.

completed the YITS questionnaire. The resulting sample size for Canada was about 30,000, much larger than for those of other countries, enabling analysis at the provincial level.

There are two class size measures available on the data. The first is self-reported by students and is their estimate of the average number of students in their language, mathematics and science classes (i.e., there is one variable for each). These variables vary within schools and reflect the different course-taking experiences of individual students. The second is the student-teacher ratio reported by principals of the schools. Both variables have advantages, but as discussed earlier, we focus on the student-teacher ratio. We use a full-time equivalent measure of teaching divided by the number of students in the school. We feel this captures the teaching input succinctly and is often a metric used by school administrators in determining and managing labour costs. Count data could provide different results since an optimal school mix might involve different combinations of part-time teachers, teaching aids and full-time faculty. Data for most schools on other than full-time teachers is thin and we feel the FTE measure is more reliable.

Time-in-term data come from the school questionnaire. These data are provided by principals and give the number of weeks in the academic year, the usual number of classes per week and the usual number of minutes per class. Days per week are not collected. Still, the data provide insight into the organization of instructional time by school administrators.

The first factor in the decompositions is the distribution of class size, the three time-in-term variables, total hours of instructional time per year and the interaction of these with the student-teacher ratio.[9] These variables provide the most complete picture available from this data on provincial variation in the allocation of teacher resources.

It should be noted that there is information from the school questionnaire on the proportion of teachers with various educational qualifications in reading, science, mathematics and education. These variables had many missing values and reduced significantly the sample available for estimation and so were not used. Unlike TIMSS, PISA did not sample intact class-rooms and so there is no teacher survey.

Student variables were chosen to reflect those factors that are best considered exogenous to the school system. For that reason, We focus on indicators of birth origin (of students and parents), single parent status, parental education and occupation, parental labour force attachment at the time of the survey, and the degree to which the student uses the language of testing at home.[10] In choosing school variables, we wanted to capture variation between the provinces in characteristics receiving a lot of attention in the academic literature and public debate. Data on other school factors come from the school questionnaire. Variables include dummy variables for population of the school community, dummies that capture the degree to which the school uses standardized tests and how student evaluations are used by school administrators and measures of teacher morale reported by school principals.

The appendix describes in more detail how we use the PISA test score data. The PISA tests and their scoring are based on the use of plausible values to quantify the abilities measured by the test items. Proficiency cutoff values were provided for the reading data to make more concrete the scoring metric used for the test results. These proficiency levels (1 to 5) reflect specific skills a student performing at that level has acquired. For details see (OECD, 2000).

V. Results

Table 1 shows the mean student-teacher ratio, total annual hours of instruction, number of weeks per year, number of classes per week and minutes of instructional time per class by province. Both student-weighted and school-weighted data are shown. We see that on average, Alberta classes are bigger (as measured by the student-teacher ratio) and that students receive more annual instructional time than in other provinces. The average numbers however, mask some important features of the class size distribution and the organization of teaching time. Table 2 shows the distribution of class size in size categories by province. It can be seen that there are very few schools with extremely small classes (less than 10) and that Alberta's proportion of these is comparable to other provinces. Alberta has a much smaller share of schools in the 10 to 19 size category and a much larger share of schools in the 20-29 size category. Alberta also has a large share of schools in the greater than 30 category but this share is comparable to that in some other provinces such as Nova Scotia, Manitoba and Québec. Table 3 shows the modal weeks per year, classes per week and minutes per class as well as the percentage of schools below the modal value. Most Canadian schools have 40 or fewer weeks in their academic year though there are more Alberta schools at the modal

---

[9] Total hours per year of instruction is equivalent to the interaction of the weeks/year, classes/week and minutes/class variables.
[10] At the time this analysis was done, variables from the YITS parents' questionnaire were not available.

value of 40 than other provinces. The most notable differences in the organization of instructional time are in the number and duration of classes. In the rest of Canada, there are fewer classes per week: 84 percent of schools have 30 or fewer classes. In Alberta, the modal number of classes per week is 40. As a consequence, typical class duration is 75 minutes in the rest of Canada whereas in Alberta there is much greater variation. 17 percent of schools have classes of 40 minutes and about 2/3 of schools have classes less than 1 hour in length. In summary, while Alberta has a larger student teacher ratio than other provinces on average, students there receive more total time in instruction broken up in more frequent, shorter classes per week.

Figures 1 to 3, show the differences between Alberta and each of the other provinces in the achievement distributions for reading, mathematics and science respectively. The vertical bars represent an indicator function that takes non-zero values at points where the two densities are statistically significantly different at the 95 percent level. The two densities are significantly different at a given point if their confidence intervals at that point do not overlap. The confidence intervals are computed using the Balanced Repeated Replication (BRR) weights provided with the data. The patterns are consistent with those when examining only the mean performance. The eastern provinces differ the most from Alberta as noted by the larger test score region in which the densities significantly differ.

An advantage of the kernel density estimator and the DFL decompositions is that they permit an easy, graphical depiction of the impact of various factors on the observed differences in test scores. We present and discuss the graphical results first. Throughout this section, the term “class size” refers to the student-teacher ratio.

V.1 The contribution of differences in student-teacher-ratios and time-in-term

Figure 4 shows the effect of fixing the distribution of class size and time-in-term at the Alberta level on the reading achievement distributions of the Atlantic provinces. The graphs show noticeable improvements in the distribution for Newfoundland-Labradour and New Brunswick. In Newfoundland, there is a spike in the density around level 4 proficiency. In New Brunswick, more students would perform above the level 5 proficiency. In Prince Edward Island a larger number of students would perform between the level 1 and 2 proficiency and between the level 3 and 4 proficiency. In Nova Scotia there would be virtually no change.

Looking more closely at the proportion of students below level 1 proficiency, we see that there would be little improvement in New Brunswick but some improvement in Nova Scotia. Thus the contribution of differences in student-teacher ratios and time-in-term is not constant throughout the achievement distribution. The Alberta student-teacher ratio and time-in-term distribution would disadvantage the poorest performing students in reading in these provinces.

Similar results are observed for the mathematics assessment. The OECD provided no proficiency intervals for the mathematics or science assessment results unlike the reading assessment. Figure 5 shows the achievement distribution in the Atlantic provinces if the student-teacher ratios and time-in-term were distributed as they are in Alberta. The vertical line indicates the international average test score (500). In Newfoundland more students would perform above the international average. In Prince Edward Island, the reverse is true. There would be virtually no change in the achievement distribution of Nova Scotia and the achievement distribution in New Brunswick would be shifted to the right, except for the lower tail, where a similar proportion of students would perform.

Figure 6 shows the results for the science assessment. For Newfoundland-Labradour, more students would be performing below the international average. For Prince Edward Island, more students would perform at or just below the international average, but at the expense of higher test scores, not lower ones. Nova Scotia would see a small improvement above the international mean at the expense of lower scores. New Brunswick shows a clear benefit—the distribution is shifted almost entirely to the right.

The graphical results suggest that if Alberta's distribution of class size and time-in-term prevailed in the Atlantic provinces, the resulting distribution of test scores would depend on the province and assessment being considered. In most cases, some students would benefit while others would not. This is probably most clear in the case of New Brunswick where students in the lowest reading proficiency would not benefit.

V.2 The contribution of differences in other-school-factors

Student-teacher ratios and time-in-term are just two of the school factors considered in this study. When we further fix the distribution of other school factors to reflect their distribution in Alberta, we see little difference in the counterfactual distributions for Newfoundland and Prince Edward Island, a slight reduction in the proportion of students performing between level 1 and level 3 proficiency in Nova Scotia, and a very slight increase in the proportion of students performing below level 1 proficiency in New Brunswick.

When we look at the mathematics assessment, we see little change for New Brunswick and Nova Scotia and more evenly distributed test scores in Newfoundland-Labradour. Here, many more students would be performing below the international average. Lastly, Figure 9 shows the effect of further fixing the distribution of student-teacher ratios and time-in-term for the science assessment. Here, we see an improvement in the proportion of students performing above the international mean for Newfoundland-Labradour and fewer students performing in the lower tail in Nova Scotia.

V.3 The contribution of differences in student-background-factors

Differences in the distribution of student population characteristics seem to have their impact at the lower tail of the distribution. Figure 10 shows the effect on differences in reading of fixing student background factors so that they are distributed as in Alberta. We see that these differences tend to contribute to the gaps with Alberta. In Newfoundland-Labradour, more students are performing at the level 4 proficiency, in Prince Edward Island, more people are performing at the level 3 proficiency. There is little impact for Nova Scotia. In New Burnswick, more students are performing near level 5 proficiency and, perhaps more importantly, fewer students are performing below level 1 proficiency. Similar patterns are observed for the mathematics assessment (Figure 11.) Lastly, the same is observed in the science assessment. Fixing student background factors to be distributed as in Alberta could mean that more students perform at the international mean score in Newfoundland-Labradour and Prince Edward Island. Fewer students would perform below the international mean in Nova Scotia and in New Brunswick.

V.4 Decomposition of selected statistics: Reading distribution

The estimated densities can be used to compute various statistics including mean performance. We present decompositions of selected achievement statistics in tables 4 to 6 for the Atlantic provinces. There were no statistically significant differences between provinces in measures of achievement inequality (such as the ratio of the 80^{th} to 20^{th} percentile or the 90^{th} to 10^{th} percentile). Therefore, we do not provide decompositions of measures of inequality as was done in by DFL. We decompose differences in the mean and standard deviations of the achievement distributions for all assessments. For the Reading assessment, we also decompose differences in the proportions of students scoring within the various proficiency intervals. Corresponding intervals are not available for the Mathematics and Science assessments. For these, we decompose differences in deciles.

For each province, the "Actual" row refers to statistics from the actual density estimated for the province. The next rows refer to the three counterfactual density functions. The first is the effect of fixing student-teacher ratios and time-in-term to have their Alberta distribution, the next further fixes other student factors to be distributed as in Alberta and the last refers to the case where student background factors are distributed as in Alberta.

Fixing the student-teacher ratios and time-in-term to be distributed as in Alberta, average performance in New Brunswick would increase to be equal to that of Alberta at 550. We see that this comes primarily from a reduction in the proportion of students performing between level 1 and 3 proficiency and an increase in the proportion performing above level 5 proficiency. Holding other school factors at their Alberta distribution results in an increase in the proportion of students performing at the lower proficiency levels and a decrease in the proportions performing above level 4. The result is a mean that is lower than the class size only case but higher than the original mean. New Brunswick would do better on average with the class size and time-in-term distribution and these differences contributed most it seems to differences between New Brunswick and Alberta in mean achievement. The relative contribution, however, of student-background and school factors to the total achievement gap differs across the achievement distribution. Class size and time-in-term matter more in the upper part of the distribution. Student background differences matter more at the low end of the achievement distribution.

These patterns, though, are not true throughout the Atlantic provinces, as the graphical results suggest. In Newfoundland-Labrador, the spike around the level proficiency observed in figure 4 translates into a 20 percentage point increase in the proportion of students performing between levels 3 and 4. This increase comes both from a decrease in the number performing above level 5 as well as the numbers performing in the lowest proficiency levels. In Nova Scotia, differences in the distribution of class size and time-in-term actually contribute to differences in the lower tail of the distribution; in their absence, fewer students would perform below level 2 proficiency. The result is a small increase in mean performance, suggesting that class size and time-in-term differences contribute less to the mean achievement gap between Nova Scotia and Alberta than they do to that between New Brunswick and Alberta. In Newfoundland-Labradour, differences in the distribution of class size and time-in-term contribute to differences in mean achievement as well, but to a lesser extent than for New Brunswick; in the absence of these differences, mean achievement in Newfoundland-Labradour would be higher. This comes from a larger proportion of students performing at the level 3 to 4 proficiency range. Thus these differences contribute most to the gap here. They *reduce* the gap in the number of students performing in the lower proficiency ranges in Newfoundland-Labradour. In their absence, more students would perform between level 1 and 2 proficiency. Interestingly, student background factors work in the same part of the achievement distribution for Newfoundland.

Tables 5 shows the statistics for the mathematics assessment. As mentioned, there were no proficiency intervals defined for the mathematics nor science assessments. We decompose differences in the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles of the achievement distributions. Again, no inequality measures are computed because there were not significant differences in measures of inequality between the provinces. For Newfoundland-Labradour, fixing student-teacher ratio and time-in-term at their Alberta distribution results in higher mean and median performance which seems to be driven by better performance at the lower end of the distribution. When other school factors are also distributed as in Alberta, these gains are apparently undone. Student background differences appear to drive the math differences between Newfoundland-Labradour and Alberta. Once these are eliminated, mean and particularly median performance is up but this appears to be due to improvements at the upper end of the achievement distribution.

For Prince-Edward Island, there is little change in mean or median performance when the differences in the distribution of class-size and time-in-term are eliminated. But there are improvements in the lower half of the distribution which are masked by a lowering of performance at the upper end. When all school factors are fixed at their Alberta distribution, there are still notable improvements in the lower half of the distribution but mean and median achievement is lower. In contrast, New Brunswick would see large improvements in mean and median performance driven mostly by improvements in the upper tail of the distribution. When all school factors are fixed at their Alberta level, there is little change in mean and median performance but this is due to an improvement at the upper half of the distribution that is offset by a fall in the lower half of the distribution. For Nova Scotia, there is little impact of changing the class-size and time-in-term variables to be distributed as in Alberta. Fixing all school factors at their Alberta distribution generally improves things throughout the distribution. Differences in student background factors seem to drive differences at the lower half of the distribution.

Table 6 shows the results for the science test. We see that New Brunswick would experience very large improvements in achievement if class-size and time-in-term were distributed as in Alberta. These would be offset little by fixing all school factors at their Alberta distribution. For Nova Scotia, fixing class-size and time-in-term at their Alberta distribution would yield small improvements in mean and median performance which would come largely from the lower half of the distribution.

For both the mathematics and science assessments we tend to see varied results across provinces. Fixing the distribution of class-size and time-in-term benefits some provinces in terms of mean and median performance but this is driven by improvements in either the upper or lower half of the distribution (depending on the province and the assessment). In some cases, such as the mathematics assessment in Nova Scotia, there is no change in mean or median performance but this effect masks noticeable effects at the upper and lower half of the distribution which cancel each other out.

In results not reported here, we computed the decompositions considering first class-size, then time-in-term then the other factors. The effects of class-size were small compared to those of time-in-term but the qualitative effects of both were similar. The variation in time-in-term is greater across provinces so this is to be expected. Again, we are motivated by the idea that the teaching input is an important choice variable that can be measured not just as the full-time equivalent teaching complement by how much those teachers teach in a year and how that time is allocated. Given this we emphasize the combined results of the two measures rather than their relative effect.

V.5 The effect of the dual language systems in Nova Scotia and New Brunswick

Our analysis of the difference between Nova Scotia, New Brunswick and Alberta is complicated by the fact that these two provinces have dual language (English and French) school sectors while Alberta does not. In both provinces and for the reading and science assessments, the English sectors had better mean and median performance than the French sectors. This was not the case in mathematics where the English sector had a slightly better performance in New Brunswick and a worse mean performance in Nova Scotia. In this section we focus on the English sectors of Nova Scotia and New Brunswick. The selected statistics estimated from the counter factual distributions are compared to those in the previous sections to assess the contribution of the French sectors to the results observed there. Tables 7, 8 and 9 depict the results for these provinces for reading, math and science respectively.

For reading we see a nearly opposite effect of fixing class-size and time-in-term at its Alberta distribution for New Brunswick when we consider just the English sector. When considering both sectors (Table 4) there was a large increase in mean performance driven by a substantial improvement in all but the bottom proficiency bracket of the achievement distribution. In Table 7, mean performance is much lower with the Alberta class-size and time-in-term distribution and we see large proportions of students performing in the bottom 3 proficiency brackets. Reverse effects are observed also for other school factors and student background factors. Thus, it is important to distinguish between the French and English sectors in New Brunswick as both respond differently in our analysis to the kinds of "experiments" being conducted here. A similar result is observed for Nova Scotia.

For the mathematics assessment in New Brunswick, on the other hand, we see a similar pattern when fixing class-size and time-in-term for just the English sector—improvement in mean performance driven by larger improvements in the upper 90 – 80 percent of the distribution. The effects for the other factors is similar. The same can be said for Nova Scotia. For Science, English only and total population results seem similar for both Nova Scotia and New Brunswick.

V.6 The order of decomposition

We noted earlier that the order in which we consider the factors in the decompositions has an effect on the estimated contribution of the factors we are considering. In our decompositions thus far we have assessed differences in school characteristics before controlling for differences in student background.

We do this because the policy experiment we have in mind is what would happen if the (apparently) successful school characteristics of Alberta were adopted by other provinces *given their specific populations*. One of our primary hypotheses is that school systems organize themselves in a way that is optimal given the student population they must serve. While some may be tempted to argue that other provinces may benefit by taking on some of the characteristics Alberta's school system, we show that this may not be optimal for all students in the other province. Indeed, we saw that Alberta's school characteristics would benefit higher achieving students in New Brunswick but might hurt the lowest achieving students. Such observations raise important considerations for policy makers and for those comparing school systems across policy boundaries without considering that school systems are designed to address the features of local populations.

Nevertheless, we reverse the order in which we consider the factors as a sensitivity check. We assess differences in the student-teacher-ratio and time-in-term after fixing other school factors and student background factors to be distributed as they are in Alberta. We focus here on the reading assessment only as this was the principle focus of the 1999 PISA. Table 10 provides the decomposition of the selected reading statistics discussed in IV.4 above in reverse order.

Considering New Brunswick (the largest inter-provincial achievement difference) we see that fixing class-size and time-in-term to be distributed as in Alberta increases mean performance by reducing the proportion of students in the bottom three proficiency brackets after already fixing student background and other school factors to their Alberta distribution. Other school factors reduce mean performance in the reverse order analysis as well and as in Table 4, This is due primarily to a greater number of students performing in the lowest proficiency levels. For Nova Scotia and Newfoundland-Labradour, the reverse order results for class-size and time-in-term are similar to the primary order effects noted in Table 4.

V. 7 The Remaining Provinces

We lastly consider briefly the results of our decomposition analysis applied to the remaining provinces and their difference with Alberta in the Reading assessment. Fixing class-size and time-in-term to be distributed as they are in Alberta reduces the proportion of students performing in the bottom two proficiency cutoffs for Québec, Ontario, and Manitoba, but improves mean performance only for Québec and Manitoba. For Saskatchewan and British Columbia, mean performance would be worse and this would be driven by changes around the lower end of the achievement distribution where a much greater proportion of students would be performing in the 1 to 2 proficiency range. When all school factors are fixed at their Alberta distribution, fewer students perform in the bottom two proficiency brackets in all provinces except British Columbia.

V. Summary and Conclusion

This paper examines the contribution of various school and student-background characteristics to the differences in high school achievement distributions for 15-year olds in Canada. It focuses on differences between the province of Alberta and the Atlantic provinces as these were the largest observed in the 2000 PISA data. Our approach considers the entire distribution of test scores, not simply the mean. Our interest was on differences in the distribution of student-teacher-ratios and time-in-term, defined as the allocation of minutes per class, classes per week and weeks per academic year of instruction.

We find evidence that school factors underlie observed differences in the achievement distributions between Alberta and the Atlantic provinces. More importantly we find that removing differences in the distribution of class size and time-in-term has a number of effects depending on the province, the assessment, and the part of the distribution being considered. In some cases the difference in mean or median performance is not attributable to differences in class size and time-in-term but this lack of noticeable effect masks noticeable effects in the different parts of the distribution. In cases where differences in class-size and time-in-term contribute to mean or median differences, it is not always because the counterfactual distribution shifted entirely to the right. In many cases, the differences in class size and time-in-term *reduce* the gap with Alberta in a particular part of the distribution, as for example in the New Brunswick reading assessment. Here, our analysis suggests that eliminating the differences in class-size and time-in-term would explain the gap in mean performance but the proportion of students performing in the lowest reading proficiency level would increase. Such an observation might be due to the way in which New Brunswick schools optimally structure themselves to address the needs of the local population. It also underscores the important tradeoffs facing policy makers who seek to introduce reforms that improve average test score performance. Such reforms may not benefit all students equally and may even hurt lower performing students.

Observational studies of student achievement receive a lot of attention when they show differences in achievement between policy jurisdictions. Despite the difficulties in determining genuine treatment effects of alternative school policies using observational data, there is often pressure to try to infer what the "good" jurisdictions are doing right and to adopt those policies elsewhere. The effects illustrated in this paper show potential implications for the entire distribution of students of undertaking such a strategy. Given the covariation of policy and population variables in the real world, as well as the association of both with unobserved characteristics, we cannot know for sure whether these effects would actually take place. The best alternative is the costly undertaking of running the very experiment we are imagining in the decompositions we report. Our approach is a useful compromise. It avoids some the specification issues plaguing more ambitious attempts at estimating policy parameters and may provide some useful insight and incentive for the design of more ambitious evaluations.

Appendix

In this appendix we write out how we obtain the counterfactual density estimates for the factors we are considering. We decompose the density differences into differences in the distribution of three sets of observables, class-size and time-in-term, other school factors then student characteristics in that order. Adapting equation (2) in the text, we have

$$\begin{aligned} f_1(y; p_y = 1, p_s = 0, p_w = 1, p_x = 1) \\ &= \int_s \int_w \int_x f(y \mid s, w, x) dF(s \mid w, x, p_s = 0) dF(w \mid x, p_w = 1) dF(x \mid p_x = 1) \\ &= \int_s \int_w \int_x f(y \mid s, w, x) \psi_{s|x,w}(s, w, x) dF(s \mid w, x, p_s = 1) dF(w \mid x, p_w = 1) dF(x \mid p_x \quad) \end{aligned} \qquad \text{(A1}$$

Where

$$\begin{aligned} \psi_{s|x,w}(s, w, x) &\equiv \frac{dF(s \mid w, x, p_s = 0)}{dF(s \mid w, x, p_s = 1)} \\ &= \left(\frac{pr(p_s = 0 \mid s, w, x)}{pr(p_s = 1 \mid s, w, x)} \right) \left(\frac{pr(p_s = 1 \mid w, x)}{pr(p_s = 0 \mid w, x)} \right) \end{aligned} \qquad \text{(A2)}$$

Similarly, we obtain the counterfactual for other school factors as

$$\int_s \int_w \int_x f(y \mid s, w, x) \psi_{s|w,x}(s, w, x) \psi_{w|x}(w, x) dF(s \mid w, x, p_s = 1) dF(w \mid x, p_w = 1) dF(x \mid p_x = 1) \qquad \text{(A3)}$$

and the counterfactual for student background factors as

$$\int_s \int_w \int_x f(y \mid s, w, x) \psi_{s|w,x}(s, w, x) \psi_{w|x}(w, x) \psi_x(x) dF(s \mid w, x, p_s = 1) dF(w \mid x, p_w = 1) dF(x \mid p_x = 1) \qquad \text{(A4)}$$

The weighting functions are given by

$$\begin{aligned} \psi_{w|x}(s, w, x) &\equiv \frac{dF(w \mid x, p_w = 0)}{dF(w \mid x, p_w = 1)} \\ &= \left(\frac{pr(p_w = 0 \mid w, x)}{pr(p_w = 1 \mid w, x)} \right) \left(\frac{pr(p_w = 1 \mid x)}{pr(p_w = 0 \mid x)} \right) \end{aligned} \qquad \text{(A5)}$$

$$\psi_{w|x}(s,w,x) \equiv \frac{dF(x \mid p_x = 0)}{dF(x \mid p_x = 1)}$$
$$= \left(\frac{pr(p_x = 0 \mid x)}{pr(p_x = 1 \mid x)}\right)\left(\frac{pr(p_x = 1)}{pr(p_x = 0)}\right)$$

With the weighting functions in hand we obtain the kernel estimates of each counterfactual by weighting all the observations in province 1 according to the following:

$\theta' = \theta\psi(s,w,x)$	For class-size and time-in-term
$\theta'' = \theta'\psi(w,x)$	For other school factors
$\theta''' = \theta''\psi(x)$	For student background factors

Finally, simplifying the notation $\widehat{f}(y; p_y = i, p_s = i, p_x = i, p_w = i)$ to $\widehat{f}_{iiii}$, we decompose the differences in densities as follows:

$$\widehat{f}_{1111} - \widehat{f}_{0000} = \widehat{f}_{1111} - \widehat{f}_{1011} + \widehat{f}_{1011} - \widehat{f}_{1001} + \widehat{f}_{1001} - \vec{f}_{1000} + \widehat{f}_{1000} - \widehat{f}_{0000} \quad (7)$$

The differences to the right of the "equals" sign represent in order, the contribution of differences in student-teacher ratio and time-in-term, the contribution of other school factors, the contribution of student background factors and a residual. The order of decomposition could potentially be important so we also decompose the difference in densities in reverse order as a point of comparison.

The clustered nature of the data does not present any direct issues for estimation as it would in the linear regression context. The clustering suggests that outcomes are correlated within schools, which, in the regression context, violates an assumption of the classical linear regression model. This is a widely cited reason for using estimation approaches like the HLM to estimate regression parameters. Kernel estimates of the density function do not require assumptions about the independence of observations. Inference, however, is affected by the correlation. As indicated above, variance estimation for nonparametric regression and density estimates is an open discussion in the literature. Replication methods like the bootstrap or the jackknife are often recommended.[11] Replication methods are advantageous when using complex survey data (like the Canadian YITS/PISA data) if the sample selection processes are applied in producing the replicate samples and survey weighs are recalculated accordingly. For PISA, balanced repeated replication was used to provide 80 replicate samples for variance estimation. Each sample is represented by a unique weight and these weights can be used to calculate the sampling variation of a statistic estimated with from the data.

The dependent variable in this paper is the reading test results. There are actually five variables for each student that reflect their performance on the test. These "plausible values" are a means of estimating aggregate population statistics (such as mean performance) that do not suffer from biases inherent in other estimation methods, particularly with tests of relatively few items.[12] Aggregate statistics can be estimated with any one set of plausible values. It is recommended, however, to use all five values. In the case of the density estimates used here, this means that the reported density (in the notation of equation (7)) is

$$\widehat{f}_{iiii} = \frac{1}{J}\sum_{j}\widehat{f}_{iiii,j} \qquad \text{(A6)}$$

where j indexes the J plausible values. The use of plausible values introduces another source of variation associated with the process used to estimate them. If v_j is the sampling variation of $\widehat{f}_{iiii,j}$, then

$$v = \frac{1}{J}\sum_{j=1}^{J} v_j + \left(1+\frac{1}{J}\right)\left(\frac{1}{J-1}\right)\sum_{j=1}^{J}(\widehat{f}_{iiii,j} - \widehat{f}_{iiii})^2 . \qquad \text{(A7)}$$

As mentioned previously, in this analysis, the sampling variances v_j are obtained using the balanced repeated replication (BRR) weights provided with the PISA data.

[11] Donald, Green and Paarsch (2000) provide an alternative nonparametric estimator of the cumulative distribution function (CDF) that is based on the calculations used to obtain hazard rates. Their approach allows specification of standard errors as well as calculations of marginal effects.
[12] For a discussion of plausible values see Mislevey (1991). For more general discussions in the context of the PISA achievement data see OECD (2000).

References

Angrist, J. and Lavy, V. (1999). Using Maimonides" Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics 114 (2)* , 533–575.

Bedard, K., Brown jr., W. O. & Helland, E. (1999). School Size and the Distribution of Test Scores. Claremont College working papers in Economics no. 1999–11.

Blundell, R., & Duncan. A. (1998). Kernel Regression in Empirical Microeconomics. *The Journal of Human Resource*s 33 (1) , 62–87.

Card, D. & Krueger. A. B. (1992). Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *The Journal of Political Economy 100 (1)* , 1–40.

Cook, M. D., & Evans, W. (2000). Families or Schools? Explaining the Convergence in White and Black Academic Performance. *Journal of Labor Economics 18 (4)* , 729–754.

Corak, M. & Lauzon, D. (2002). Provincial Differences in High School Achievement, for Whom Do Schools Matter? Paper presented at Symposium 2002: Using Survey Data for Social and Economic Research. Ottawa. November.

Denham, C., & Leiberman, A. (eds.). (1980). *Time to Learn.* (California Commission on Teacher Preparation and Licensing).

DiNardo, J. & Tobias, J.L. (2001). Nonparametric Density and Regression Estimation. *Journal of Economic Perspectives 15 (4)* , 11–28.

DiNardo, J., Fortin, N.M. & Lemieux, T. (1996). Labour Market Institutions and the Distribution of Wages 1973–1992: A Semiparametric Approach. *Econometrica. 64 (5)* , 1001–1044.

Doiron, D. & Riddell, W. C. (1994). The Impact of Unionization on Male-Female Earnings Differences in Canada. *Journal of Human Resources 29 (2)*, 504-534

Donald, S.G., Green, D. A., & Paarsch, H. J. (2000). Differences in Wage Distributions Between Canada and the United States: An Application of a Flexible Estimator of Distribution Functions in the Presence of *Covariates. Review of Economic Studies 67 (4)* , 609–633.

Economic Council of Canada. (1992). *A Lot to Learn: Education and Training in Canada, a Statement*. Ottawa.

Ehrenberg, R.G., Brewer, D.J. , Gamoran, A. & Willms, J.D. (2001). Class Size and Student Achievement. *Psychological Science in the Public Interest. 2 (1)* , 1–30.

Eide, E. & Showalter, M.H. (1998). The Effect of School Quality on Student Performance: A Quantile Regression Approach. *Economic Letters. 58 (3)* , 345–350.

Fairlie, R. 1999. *The absence of the African-American owned business: An analysis of the dynamics of self-employment.* Journal of Labor Economics, 17 (1), pp. 80-108.

Hanushek, E. (1979). Conceptual and Empirical Issues in the Estimation of Educational Production Functions. *Journal of Human Resources. 14 (3)* , 351–388.

Hanushek, E. (1986). The Economics of Schooling: Production and Efficiency in the Public Schools. *Journal of Economic Literature. 24 (3)* , 1141–1178.

Hanushek, E. (1996). School Resources and Student Performance. In *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success.* G. Burtless, (ed.). Brookings Institute. Washington D.C. 43–73.

Hanushek, E. (1998). The Evidence on Class Size. Occasional Paper 98–1. W. Allen Wallis Institute of Political Economy. University of Rochester.

Hanushek, E. & Taylor, L. (1990). Alternative Assessments of the Performance of Schools: Measurement of State Variations in Achievement. *Journal of Human Resources. 25 (2)* , 179–201.

Heckman, J. J. & Vytlacil, E. (2005). Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica 73 (3), 669 -738*

Hedges, L.V., Laine, R., & Greenwald. R. (1994). Does Money Matter? A Meta- analysis of Studies of the Effects of Differential School Inputs on Student Outcomes. *Educational Researcher. 23 (3)* , 5–14.

Hoxby, C.M. (2000). The Effects of Class Size on Student Achievement: New Evidence from Population Variation. *The Quarterly Journal of Econo*mics. *115 (4)* , 1239–1285.

Krueger, A. (1997). Experimental Estimates of Education Production Function. NBER working paper 6051.

Lazear, E. (1999). Educational Production. NBER working paper 7349.

Levin, J. (2001). For Whom the Reductions Count: A Quantile Regression Analysis of Class Size and Peer Effects on Scholastic Achievement. *Empirical Economics. 26 (1)* , 221–246.

Loeb, S. & Bound, J. (1996). The Effect of Measured School Inputs on Academic Achievement: Evidence from the 1920s, 1930s and 1940s Birth Cohorts. *The Review of Economics and Statistics.* 78 (4) , 653–664.

Mislevey, R. J. (1991). Randomization-based Inferences about Latent Variables from Complex Samples *Psychometrika. 56* , 177–196.

Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labour Markets. *International Economic Review. 14 (3)* , 693–709.

OECD. (2002). PISA Technical Report. OECD Publications. Paris.

Park, B.U. & Turloch, B.A. (1992). Practical Performance of Several Data Driven Bandwidth Selectors. *Computational Statistics. 7* , 251–270.

Raudenbush, S.W. & Bryk, A.S. (1986). A Hierarchical Model for Studying School Effects. *Sociology of Education. 59 (1)* , 1–17.

Raudenbush, S. W. & Willms, J.D. (1995). The Estimation of School Effects. *Journal of Educational and Behavioral Statistics. 20 (4)* , 307–335.

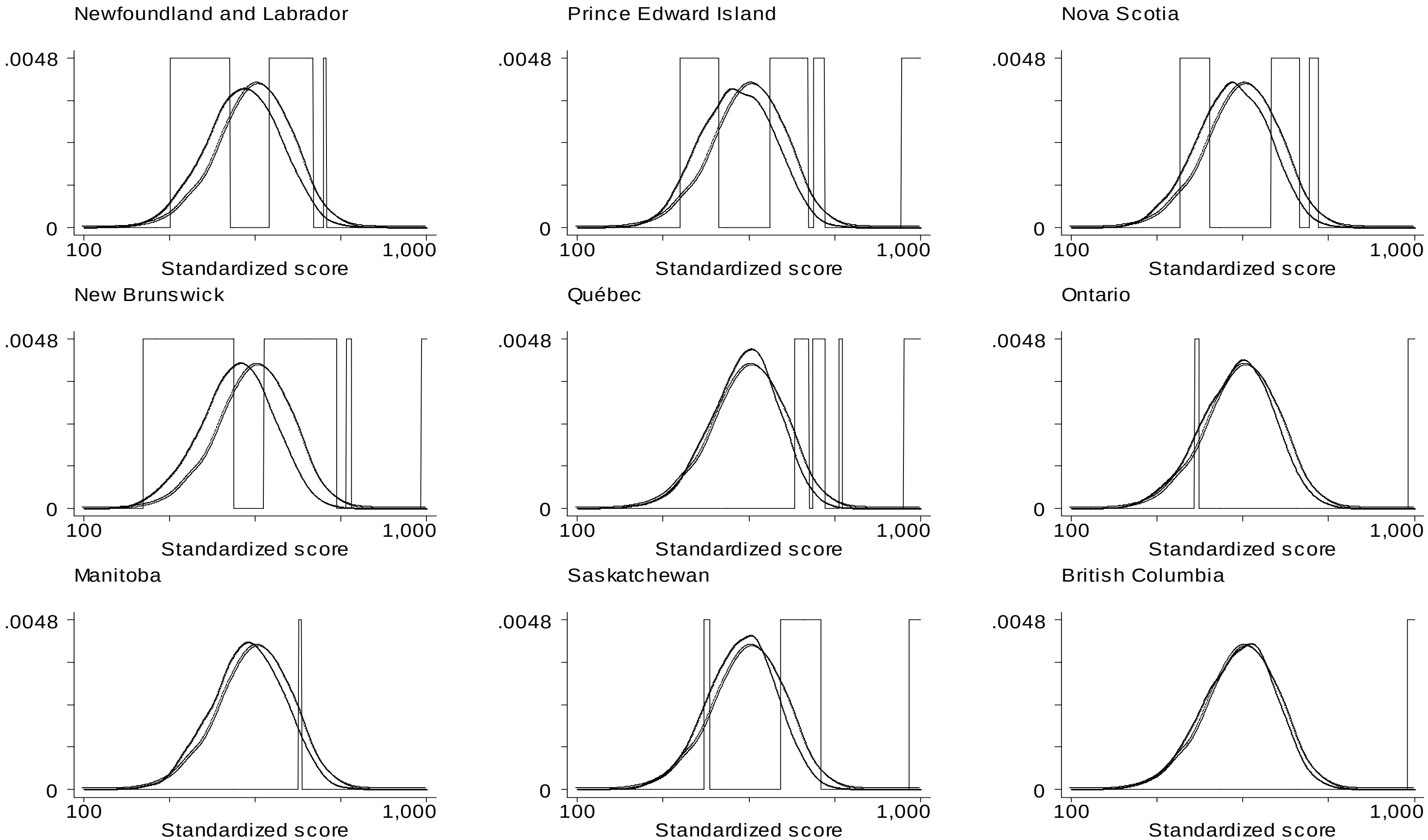
Sheather, S. & Jones, M. (1991). A Reliable Data-based Bandwidth Selection Method for Kernel Density Estimation. *Journal of the Royal Statistical Society. Series B, 53 (3)* , 683–690.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. London. Chapman and Hall.

Willms, J.D. & Raudenbush, S. (1989). A Longitudinal Hierarchical Linear Model for Estimating School Effects and Their Stability. *Journal of Educational Measurement. 26 (3)* , 209–232.

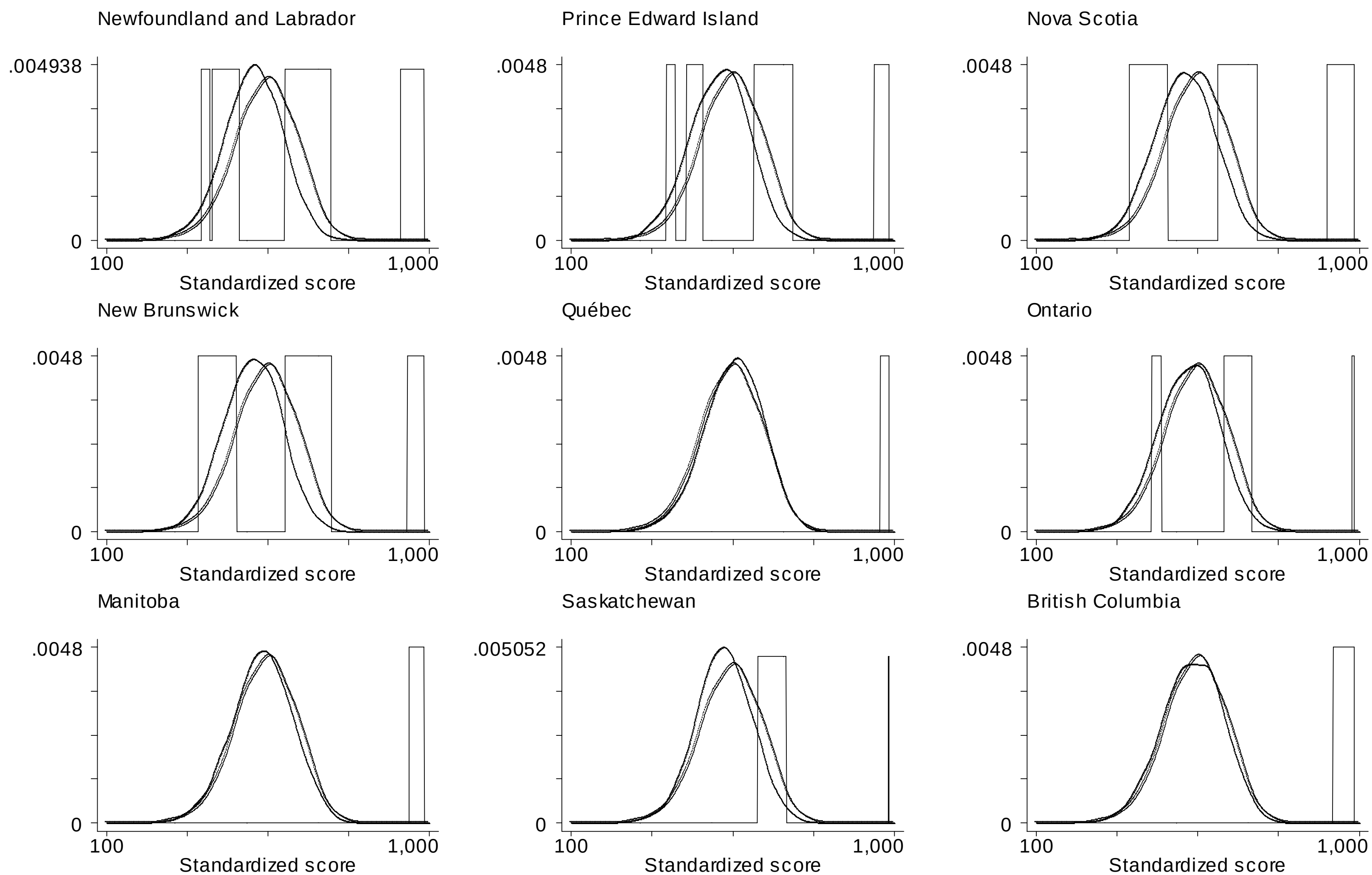
Yatchew, A. (1998). Nonparametric Regression Techniques in Economics. *Journal of Economic Literature. 36 (2)* , 669–721.

Figure 1
Provincial differences relative to Alberta (thick line) in achievement distributions, Reading



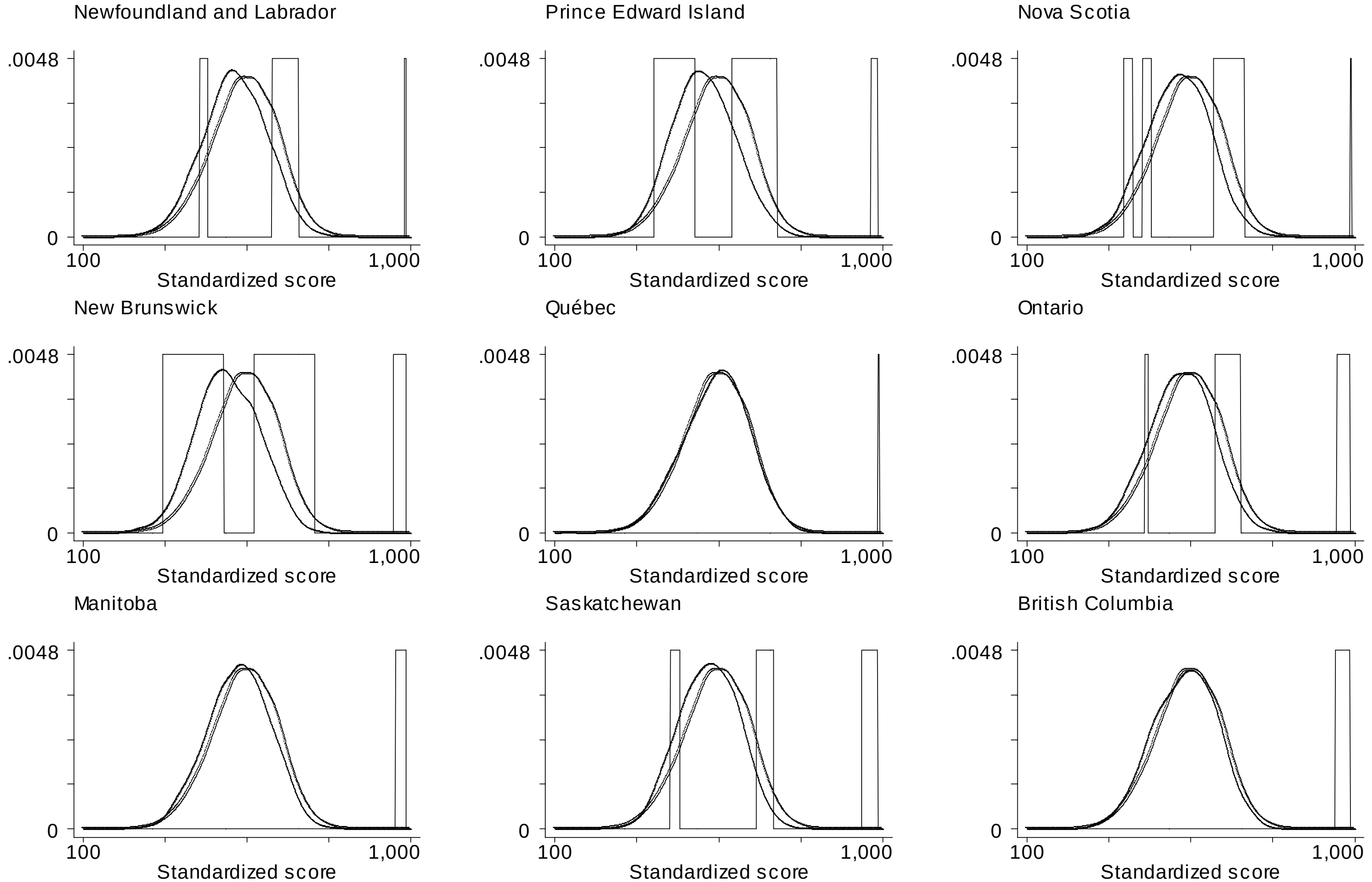
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Figure 2
Provincial differences relative to Alberta (thick line) in achievement distributions, Mathematics



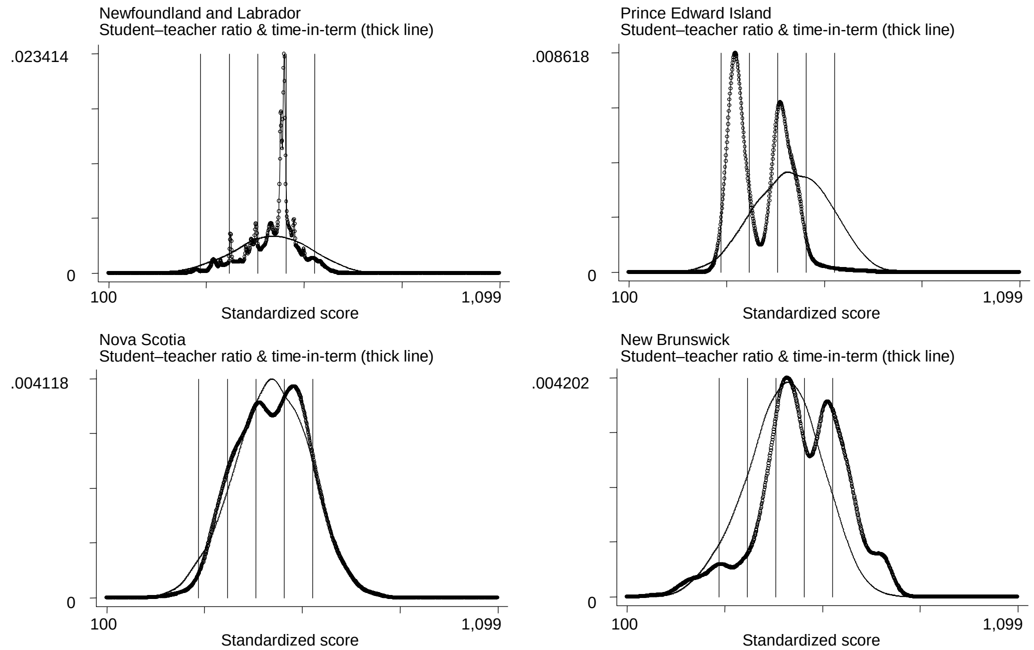
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Figure 3
Provincial differences relative to Alberta (thick line) in achievement distributions, Science



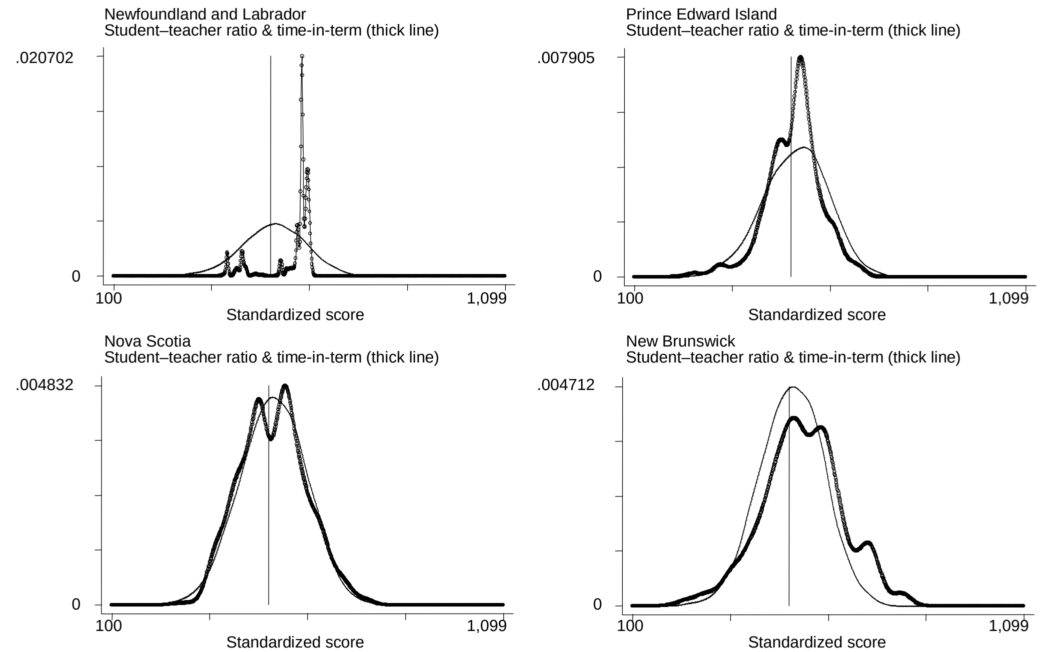
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Figure 4
The contribution of student–teacher ratios and time-in-term: Reading assessment, Atlantic provinces[a]



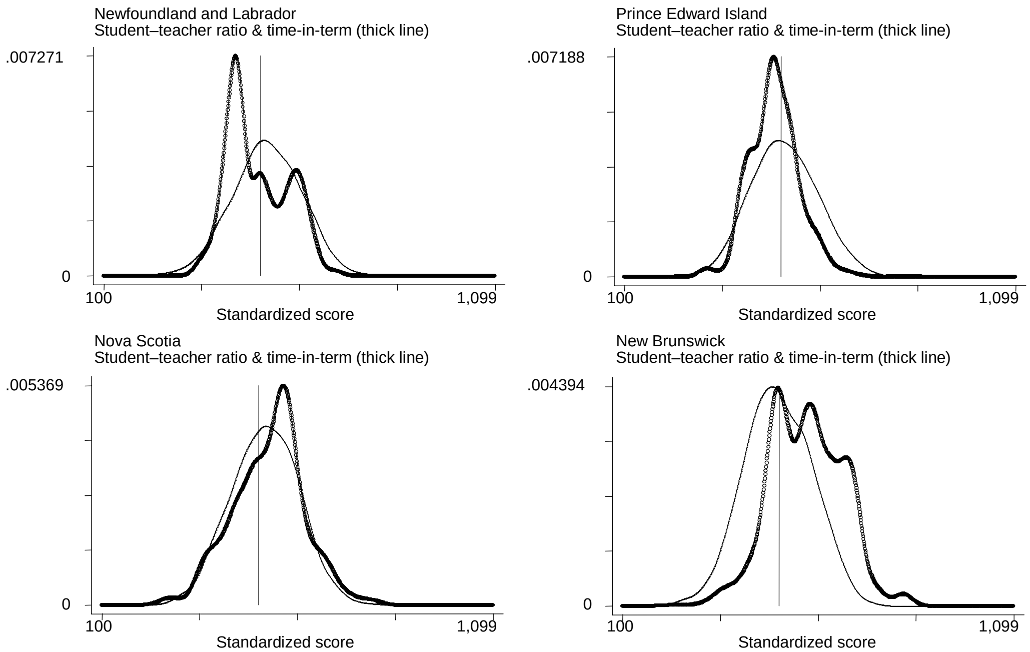
[a] Vertical lines depict proficiency level cutoffs.
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Figure 5
The contribution of student–teacher ratios and time-in-term: Mathematics assessment, Atlantic provinces[a]



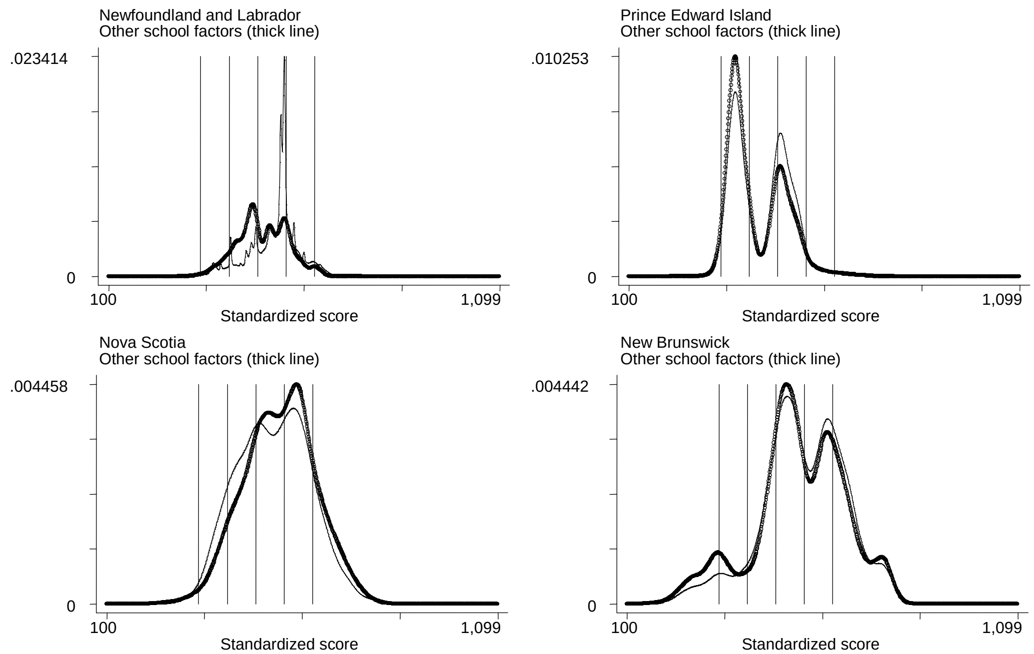
[a] Vertical line depicts the mean PISA result for the province depicted in the panel.
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Figure 6
The contribution of student–teacher ratios and time-in-term: Science assessment, Atlantic provinces[a]



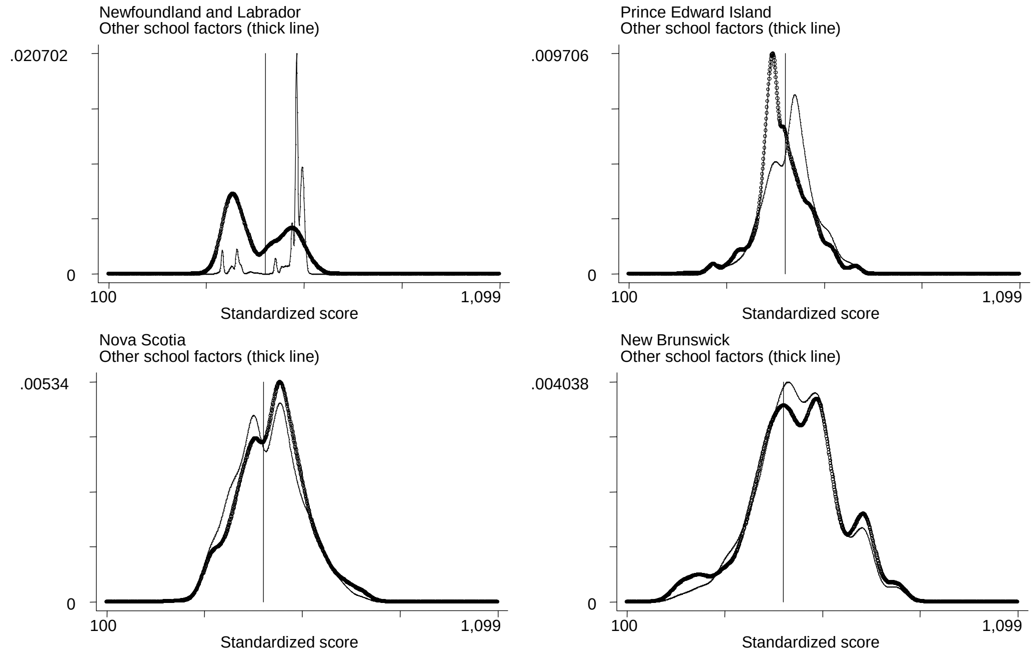
[a] Vertical line depicts the mean PISA result for the province depicted in the panel.
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Figure 7
The contribution of other school factors: Reading assessment, Atlantic provinces[a]



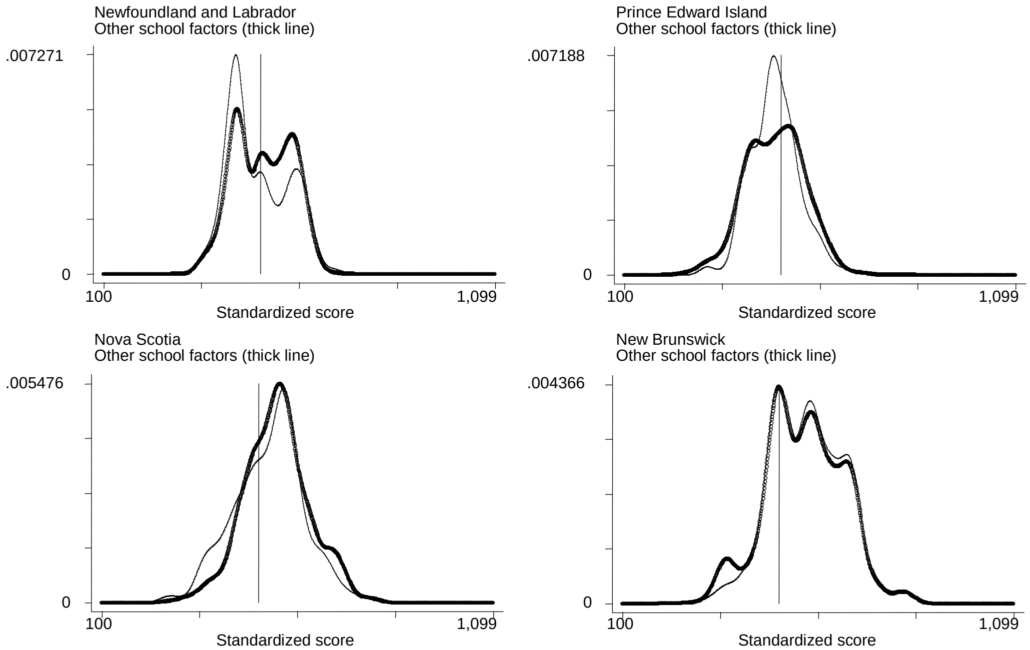
[a] Vertical lines depict proficiency level cutoffs.
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Figure 8
The contribution of other school factors: Mathematics assessment, Atlantic provinces[a]



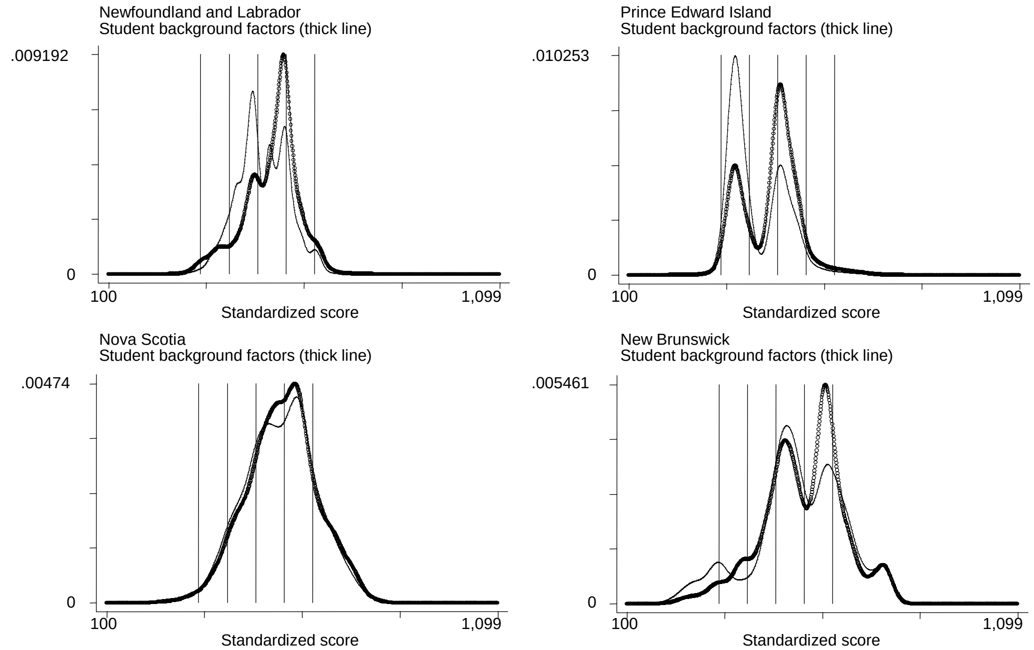
[a] Vertical line depicts the mean PISA result for the province depicted in the panel.
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Figure 9
The contribution of other school factors: Science assessment, Atlantic provinces[a]



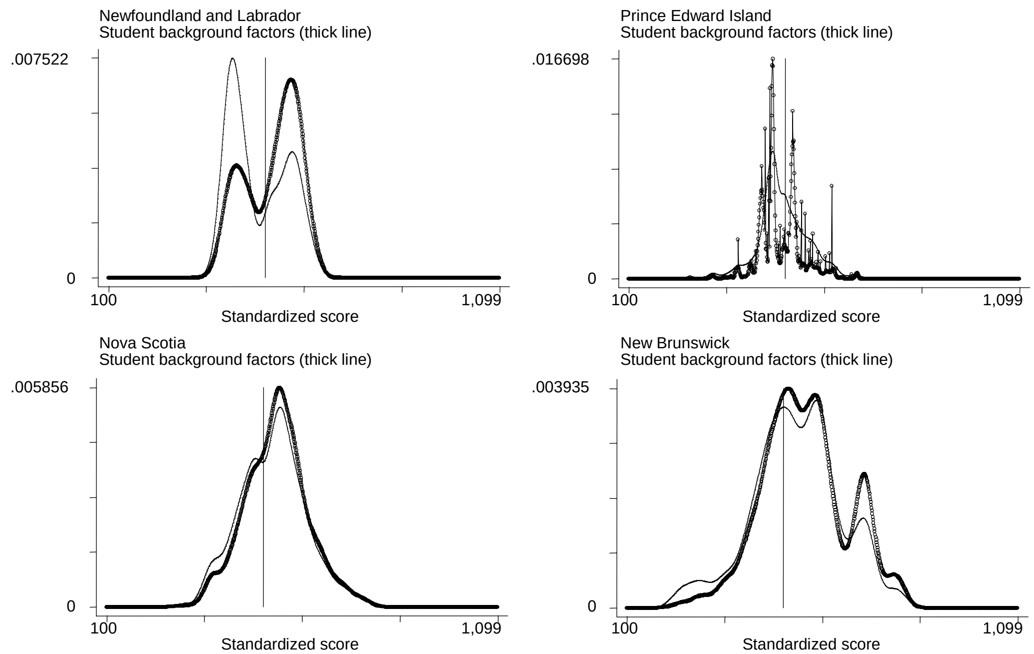
[a] Vertical line depicts the mean PISA result for the province depicted in the panel.
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Figure 10
The contribution of student background factors: Reading assessment, Atlantic provinces[a]



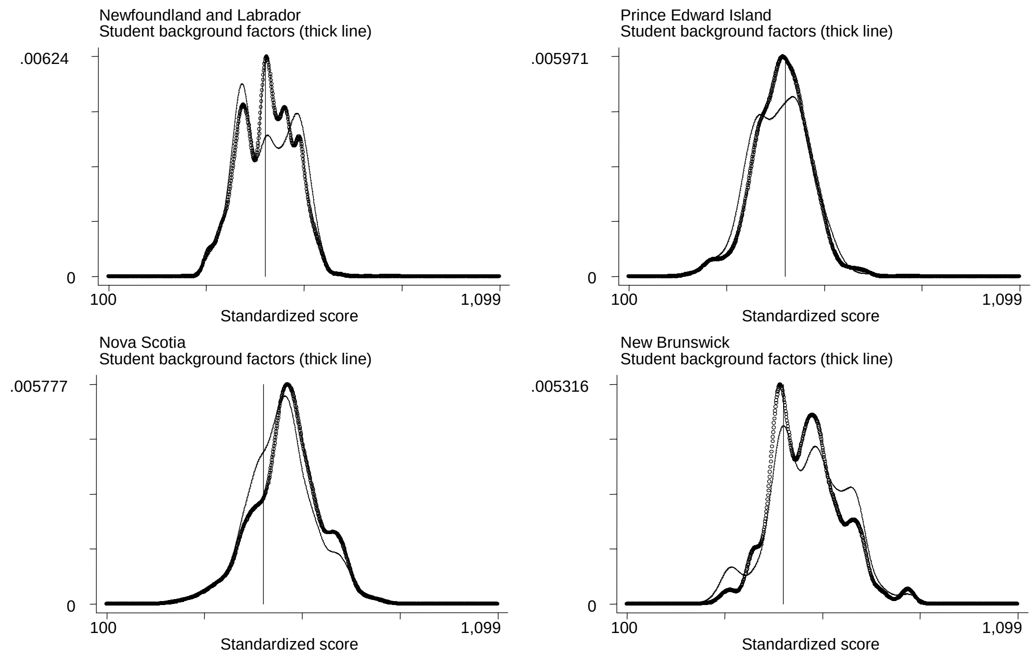
[a] Vertical lines depict proficiency level cutoffs.
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Figure 11
The contribution of student background factors: Mathematics assessment, Atlantic provinces[a]



[a] Vertical line depicts the mean PISA result for the province depicted in the panel.
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Figure 12
The contribution of student background factors: Science assessment, Atlantic provinces[a]



[a] Vertical line depicts the mean PISA result for the province depicted in the panel.
Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Table 1
Average student–teacher ratio, total instructional hours per academic year, weeks per year, classes per week and minutes per class, by province

	Student–teacher ratio	Total instructional hours per academic year	Weeks per academic year	Classes per week	Minutes per class
Student Weighted[1]					
Newfoundland and Labrador	15.9	893.9	36.8	26.0	58.7
Prince Edward Island	18.4	986.1	39.3	21.9	71.9
Nova Scotia	16.6	950.5	38.3	28.1	58.4
New Brunswick	17.7	955.8	38.8	22.7	66.3
Quebec	17.3	970.3	37.2	24.0	68.1
Ontario	16.1	937.1	38.9	19.9	74.4
Manitoba	16.6	1027.1	38.9	26.8	64.7
Saskatchewan	17.6	941.7	38.2	26.7	57.5
Alberta	19.3	1054.0	39.5	25.2	68.8
British Columbia	17.4	975.3	39.5	20.3	76.3
School Weighted					
Newfoundland and Labrador	14.8	917.9	37.3	27.8	57.7
Prince Edward Island	17.1	975.4	39.6	27.0	59.3
Nova Scotia	15.5	941.9	38.3	30.9	51.6
New Brunswick	16.8	954.3	38.9	23.5	64.1
Quebec	17.1	976.5	37.1	25.2	65.5
Ontario	14.9	939.9	38.8	20.2	73.7
Manitoba	15.8	1025.4	39.3	30.7	58.5
Saskatchewan	16.1	955.7	38.3	29.5	53.1
Alberta	18.7	1039.7	39.4	31.0	55.1
British Columbia	17.3	951.9	39.4	21.4	71.0

[1]Average school characteristics of the student population.

Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Table 2
Proportion of schools in selected size categories, by province

	Less than 10	10-19	20-29	30 or more
Newfoundland and Labrador	0.051	0.753	0.000	0.135
Prince Edward Island	0.000	0.593	0.000	0.259
Nova Scotia	0.026	0.734	0.018	0.174
New Brunswick	0.000	0.738	0.052	0.131
Quebec	0.014	0.621	0.116	0.189
Ontario	0.047	0.777	0.013	0.137
Manitoba	0.000	0.750	0.019	0.197
Saskatchewan	0.000	0.829	0.045	0.093
Alberta	0.032	0.410	0.197	0.240
British Columbia	0.000	0.659	0.079	0.116

Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Table 3
Modal and other selected values for organization of instructional time
Alberta and the rest of Canada

		Alberta	Rest of Canada
Weeks per year	Mode	40	40
	Proportion at mode	0.506	0.383
	Proportion below mode	0.367	0.523
Classes per week	Mode	40	20
	Proportion at mode	0.307	0.406
	Proportion below mode	0.658	0.065
	Proportion at or below 30	0.47	0.84
Minutes per class	Mode	40	75
	Proportion at mode	0.172	0.322
	Proportion below mode	0	0.555
	Proportion less than 60	0.667	0.284

Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Table 4
Decomposition of selected statistics in reading achievement: Atlantic provinces

Province/ Distribution	Mean	Standard deviation	Proficiency levels					
			Below 1	1 to 2	2 to 3	3 to 4	4 to 5	5 and above
Alberta								
Actual	550.40	98.10	0.019	0.062	0.15	0.262	0.278	0.229
Newfoundland and Labrador								
Actual	516.77	99.66	0.040	0.10	0.21	0.28	0.23	0.14
Student–teacher ratio and time-in-term	522.83	64.82	0.01	0.06	0.17	0.49	0.23	0.04
Other school	494.86	64.68	0	0.08	0.36	0.35	0.17	0.02
Student background	517.89	68.61	0.01	0.07	0.18	0.42	0.28	0.04
Prince Edward Island								
Actual	517.46	95.95	0.03	0.11	0.22	0.28	0.24	0.14
Student–teacher ratio and time-in-term	439.64	71.73	0.02	0.43	0.19	0.32	0.03	0.01
Other school	427.23	72.38	0.03	0.52	0.17	0.24	0.03	0.01
Student background	463.60	72.36	0.02	0.27	0.24	0.41	0.05	0.02
Nova Scotia (both sectors)								
Actual	521.17	95.74	0.03	0.09	0.21	0.29	0.24	0.14
Student–teacher ratio and time-in-term	526.02	93.59	0.01	0.10	0.22	0.26	0.27	0.15
Other school	543.59	91.59	0.01	0.06	0.18	0.27	0.29	0.18
Student background	548.92	90.64	0.01	0.05	0.16	0.28	0.31	0.19
New Brunswick (both sectors)								
Actual	501.15	97.49	0.05	0.12	0.23	0.29	0.21	0.10
Student–teacher ratio and time-in-term	549.99	110.35	0.05	0.05	0.14	0.28	0.23	0.26
Other school	539.70	118.81	0.07	0.05	0.13	0.29	0.22	0.24
Student background	553.08	104.27	0.03	0.06	0.14	0.25	0.29	0.23

Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Table 5
Decomposition of selected statistics in mathematics achievement, Atlantic provinces

Province/Distribution	Mean	Standard deviation	Selected Percentiles				
			10th	25th	50th (median)	75th	90th
Alberta							
Actual	546.97	86.94	433.96	488.45	549.04	607.69	657.08
Newfoundland and Labrador							
Actual	509.16	81.99	403.27	454.49	510.20	565.00	612.07
Student–teacher ratio and time-in-term	559.08	81.99	430.02	568.31	580.66	591.78	598.06
Other school	482.85	81.99	394.91	417.05	460.37	552.73	589.19
Student background	517.72	81.99	412.65	453.55	534.16	573.90	600.13
Prince Edward Island							
Actual	511.77	83.90	401.56	454.6	514.37	569.69	616.73
Student–teacher ratio and time-in-term	508.04	83.90	423.49	466.07	513.76	549.07	593.76
Other school	494.17	83.90	424.85	457.24	486.25	531.62	578.32
Student background	489.45	83.90	434.77	454.1	472.42	520.33	561.43
Nova Scotia (both sectors)							
Actual	512.60	85.40	400.61	454.34	513.19	571.27	621.82
Student–teacher ratio and time-in-term	512.67	85.40	400.61	450.92	512.03	569.06	625.50
Other school	524.75	85.40	412.78	465.42	526.82	577.74	632.00
Student background	534.91	85.40	431.11	480.96	535.76	583.23	637.60
New Brunswick (both sectors)							
Actual	506.20	82.38	398.97	450.23	507.60	563.00	609.03
Student–teacher ratio and time-in-term	534.67	82.38	401.55	468.21	533.93	600.84	673.61
Other school	533.75	82.38	391.88	462.19	534.42	606.45	690.30
Student background	554.91	82.38	422.91	481.37	548.06	620.32	710.63

Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Table 6
Decomposition of selected statistics in science achievement, Atlantic provinces

Province/Distribution	Mean	Standard deviation	Selected percentiles				
			10th	25th	50th (median)	75th	90th
Alberta							
Actual	546.32	90.49	426.77	485.81	548.68	609.81	659.19
Newfoundland and Labrador							
Actual	516.46	89.97	399.21	456.32	516.47	578.02	631.07
Student–teacher ratio and time-in-term	493.49	89.97	403.57	430.78	474.45	561.74	608.05
Other school	507.24	89.97	412.74	444.26	506.22	570.40	604.68
Student background	505.57	89.97	415.12	450.58	506.43	557.18	595.32
Prince Edward Island							
Actual	508.07	87.24	396.62	446.59	505.27	567.15	622.09
Student-teacher ratio and time-in-term	484.26	87.24	403.3	440.23	482.52	522.52	565.96
Other school	485.29	87.24	388.19	429.32	486.56	539.15	583.18
Student background	491.55	87.24	403.71	445.22	492.96	538.02	578.56
Nova Scotia (both sectors)							
Actual	515.95	88.11	399.41	455.14	517.28	577.08	626.18
Student–teacher ratio and time-in-term	527.14	88.11	400.58	465.76	535.16	584.89	639.50
Other school	547.62	88.11	443.31	491.67	546.70	598.07	660.64
Student Background	560.51	88.11	447.51	507.66	562.21	613.72	675.85
New Brunswick (both sectors)							
Actual	496.73	88.41	383.79	435.28	494.52	558.89	612.42
Student–teacher ratio and time-in-term	569.65	88.41	457.85	500.02	567.84	638.66	690.23
Other School	564.74	88.41	443.19	495.61	563.63	637.48	691.92
Student background	559.05	88.41	456.61	493.01	553.76	613.91	680.83

Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Table 7
Decomposition of selected statistics in reading achievement: English school sectors of Nova Scotia and New Brunswick

Province/Distribution	Mean	Standard deviation	Proficiency levels					
			Below 1	1 to 2	2 to 3	3 to 4	4 to 5	5 and above
Alberta								
Actual	550.40	98.10	0.02	0.06	0.15	0.26	0.28	0.23
Nova Scotia								
Actual	529.16	91.53	0.02	0.07	0.20	0.29	0.27	0.14
Student–teacher ratios and time in-term	486.16	88.11	0.05	0.16	0.26	0.28	0.20	0.05
Other school	501.35	95.45	0.04	0.14	0.24	0.28	0.20	0.10
Student background	510.34	91.61	0.03	0.13	0.20	0.32	0.23	0.10
New Brunswick								
Actual	538.04	97.42	0.02	0.07	0.17	0.26	0.28	0.19
Student–teacher ratio and time-in-term	448.40	108.06	0.08	0.41	0.19	0.11	0.12	0.09
Other school	485.62	121.37	0.07	0.30	0.16	0.14	0.18	0.16
Student background	475.99	120.4	0.07	0.33	0.16	0.12	0.16	0.15

Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Table 8
Decomposition of selected statistics in mathematics achievement, English school sectors of Nova Scotia and New Brunswick

Province/Distribution	Mean	Standard deviation	Selected percentiles				
			10th	25th	50th (median)	75th	90th
Alberta							
Actual	546.97	86.94	433.96	488.45	549.04	607.69	657.08
Nova Scotia							
Actual	512.70	85.68	400.33	454.19	513.23	571.69	622.31
Student–teacher ratio and time-in-term	510.82	85.68	398.91	448.81	509.91	567.20	623.51
Other school	523.32	85.68	410.51	463.50	525.59	576.78	630.43
Student background	534.76	85.68	431.26	481.92	535.91	582.39	635.87
New Brunswick							
Actual	504.84	83.37	397.29	447.51	505.91	562.37	609.64
Student–teacher ratio and time-in-term	526.17	83.37	397.26	476.59	538.74	587.55	627.10
Other school	523.71	83.37	379.99	475.92	541.26	588.69	627.60
Student background	532.36	83.37	416.44	482.33	541.87	586.53	625.18

Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Table 9
Decomposition of selected statistics in science achievement, English school sectors of Nova Scotia and New Brunswick

Province/ Distribution	Mean	Standard deviation	Selected percentiles				
			10th	25th	50th (median)	75th	90th
Alberta							
Actual	546.32	90.49	426.77	485.81	548.68	609.81	659.19
Nova Scotia							
Actual	517.04	87.85	400.88	456.40	518.40	577.90	626.93
Student–teacher ratio and time-in-term	526.60	87.85	401.14	465.59	534.23	584.40	638.31
Other school	546.29	87.85	442.72	490.42	545.09	596.83	658.91
Student background	559.06	87.85	445.35	504.76	561.16	613.13	674.97
New Brunswick							
Actual	502.75	86.13	394.72	442.31	498.77	563.05	616.83
Student–teacher ratio and time-in-term	552.49	86.13	436.40	485.32	545.29	625.69	677.35
Other school	545.72	86.13	414.51	478.58	536.95	623.61	679.49
Student background	554.12	86.13	446.47	490.46	550.39	618.82	672.88

Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Table 10
Decomposition of selected statistics in reading achievement, Atlantic provinces (reverse order)

Province/Distribution	Mean	Standard deviation	Proficiency levels					
			Below 1	1 to 2	2 to 3	3 to 4	4 to 5	5 and above
Alberta								
Actual	550.40	98.10	0.019	0.062	0.15	0.262	0.278	0.229
Newfoundland and Labrador								
Actual	516.77	99.66	0.04	0.10	0.21	0.28	0.23	0.14
Student background	522.83	64.43	0.01	0.06	0.17	0.49	0.23	0.04
Other school	494.86	64.25	0	0.08	0.37	0.35	0.17	0.02
Student-teacher ratio and time-in-term	517.89	68.12	0.01	0.07	0.19	0.42	0.27	0.04
Prince Edward Island								
Actual	517.46	95.95	0.03	0.11	0.22	0.28	0.24	0.14
Student background	515.92	95.30	0.02	0.10	0.26	0.27	0.20	0.15
Other school	463.16	71.77	0.02	0.27	0.23	0.42	0.05	0.02
Student–teacher ratio and time-in-term	463.60	72.36	0.02	0.27	0.24	0.41	0.05	0.02
Nova Scotia (both sectors)								
Actual	521.17	95.74	0.03	0.09	0.21	0.29	0.24	0.14
Student background	533.50	100.58	0.03	0.08	0.18	0.26	0.26	0.19
Other school	535.88	105.32	0.04	0.08	0.16	0.25	0.26	0.20
Student–teacher ratio and time-in-term	548.92	90.64	0.01	0.05	0.16	0.28	0.31	0.19
New Brunswick (both sectors)								
Actual	501.15	97.49	0.05	0.12	0.23	0.29	0.21	0.10
Student background	489.17	86.65	0.01	0.13	0.41	0.23	0.12	0.10
Other school	485.50	102.11	0.05	0.19	0.3	0.19	0.16	0.11
Student–teacher ratio and time-in-term	553.08	104.27	0.03	0.06	0.14	0.25	0.29	0.23

Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).

Table 11
Decomposition of selected statistics in reading achievement, Central and Western provinces

Province/Distribution	Mean	Standard deviation	Proficiency levels					
			Below 1	1 to 2	2 to 3	3 to 4	4 to 5	5 and above
Alberta								
Actual	550.40	98.10	0.019	0.062	0.15	0.262	0.278	0.229
Quebec								
Actual	535.78	91.27	0.02	0.07	0.17	0.29	0.29	0.16
Student–teacher ratio and time-in-time	540.05	79.40	0	0.03	0.22	0.27	0.33	0.14
Other school	533.12	82.68	0	0.06	0.24	0.25	0.31	0.14
Student background	475.39	85.26	0.01	0.17	0.46	0.14	0.16	0.06
Ontario								
Actual	533.24	96.88	0.03	0.08	0.18	0.27	0.27	0.17
Student–teacher ratio and time-in-term	523.64	85.34	0	0.05	0.28	0.39	0.15	0.13
Other school	527.14	86.67	0	0.05	0.26	0.38	0.17	0.14
Student Background	525.60	71.23	0	0.03	0.21	0.50	0.17	0.09
Manitoba								
Actual	529.37	95.74	0.02	0.09	0.19	0.29	0.25	0.16
Student–teacher ratio and time-in-term	534.21	89.42	0.01	0.06	0.20	0.35	0.21	0.17
Other school	532.48	89.94	0.01	0.06	0.21	0.35	0.20	0.17
Student Background	535.37	89.99	0.01	0.06	0.20	0.35	0.21	0.17
Saskatchewan								
Actual	529.16	91.53	0.02	0.07	0.20	0.29	0.27	0.14
Student–teacher ratio and time-in-term	486.16	88.11	0.05	0.16	0.26	0.28	0.20	0.05
Other school	501.35	95.45	0.04	0.14	0.24	0.28	0.20	0.10
Student background	510.34	91.61	0.03	0.13	0.20	0.32	0.23	0.10
British Columbia								
Actual	538.04	97.42	0.02	0.07	0.17	0.26	0.28	0.19
Student–teacher ratio and time-in-term	448.40	108.06	0.08	0.41	0.19	0.11	0.12	0.09
Other school	485.62	121.37	0.07	0.30	0.16	0.14	0.18	0.16
Student background	475.99	120.40	0.07	0.33	0.16	0.12	0.16	0.15

Source: Programme for International Student Assessment (PISA) and the Longitudinal Youth in Transition Survey (YITS).