C.D. Howe Institute
Commentary

www.cdhowe.org No. 258, February 2008 ISSN 0824-8001

Education Papers

School Grades:

Identifying British Columbia's Best Schools

David Johnson

In this issue...

The study identifies which B.C. elementary schools are worthy of emulation — and those where large improvements are possible — by removing the influence of local socioeconomic factors on a school's performance.

The Study in Brief

Standardized testing is a controversial subject, particularly in British Columbia. However, as this *Commentary* argues, standardized test results can be a valuable resource as long as they are placed into the proper context. It is no surprise that students who have parents with more education or speak English as a first language do far better on standardized tests than otherwise disadvantaged students. This *Commentary* compares outcomes in British Columbia schools where students come from similar backgrounds.

Professor David Johnson's methodology, based on his ground-breaking study of Ontario schools, identifies which schools are doing better or worse than expected given the socioeconomic characteristics of their students.

In British Columbia slightly over half of the variation in tests scores is associated with variation in student backgrounds. The other half of the variation is associated with school-specific characteristics. This half of the variation allows a fair comparison of schools because the ratings in this study are constructed by comparing schools with a similar mix of students. This *Commentary* also finds strong evidence that there are schools where educators or parents appear to have influenced the composition of students actually writing the assessments and thus raised their school's rank in the Fraser Institute rankings of school quality.

Parents and educators alike should welcome a fair comparison of schools, rather than a ranking. When exceptionally strong schools are identified, then other schools can try to understand and emulate what exceptional schools are doing. The associated tables of school test scores, student socioeconomic characteristics at schools and, most importantly, the percentile ratings of schools allow a comparison of how schools actually do on the Foundation Skills Assessments compared to the scores predicted by the socioeconomic characteristics of their students. A high percentile school is an exceptionally strong school that has outperformed other similar schools. A low percentile school is a school with weak FSA results relative to other similar schools. Schools at both extremes of the percentile distribution should be further investigated.

While the critics of the Foundation Skills Assessment are right in saying that socioeconomic characteristics are a major driver of test scores, this *Commentary* provides a measure of the variation in test scores across schools with the same socioeconomic characteristics. This allows a fair comparison of schools. Standardized tests in British Columbia are a useful exercise for school comparisons as long as the data are used in an appropriate manner.

The Author of This Issue

David Johnson is Professor, Department of Economics, Wilfrid Laurier University and a Fellow-in-Residence and Education Policy Scholar, C.D. Howe Institute.

* * * * * *

C.D. Howe Institute Commentary© is a periodic analysis of, and commentary on, current public policy issues. Michael Benedict and James Fleming edited the manuscript; Diane King prepared it for publication. As with all Institute publications, the views expressed here are those of the author and do not necessarily reflect the opinions of the Institute's members or Board of Directors. Quotation with appropriate credit is permissible.

To order this publication please contact: Renouf Publishing Company Limited, 5369 Canotek Road, Ottawa, Ontario K1J 9J3; or the C.D. Howe Institute, 67 Yonge St., Suite 300, Toronto, Ontario M5E 1J8. The full text of this publication is also available on the Institute's website at www.cdhowe.org.

$12.00; ISBN 0-88806-732-1
ISSN 0824-8001 (print); ISSN 1703-0765 (online)

How can parents, teachers, taxpayers and school administrators know how well elementary schools in British Columbia perform in imparting knowledge and skills? Is there more information than parents can glean from occasional parent-teacher nights? Is there more information than is contained in the traditional school performance measures and rankings?

This study is designed to compare outcomes in British Columbia schools where students come from similar backgrounds. This analysis identifies schools worthy of emulation and praise, as well as schools where the evidence suggests large improvement is possible. Parents, teachers, taxpayers and school administrators should be able to use this new information to increase their understanding of how well specific schools are doing. This analysis will help all participants in the British Columbia education system to better understand what can and cannot be learned from the province's annual Grades 4 and 7 Foundation Skills Assessment (FSA) process — and how use of this assessment could be improved.

The FSA test results in numeracy, reading and writing offer one way of measuring school performance. British Columbia conducts the FSA for students in Grades 4 and 7 at all schools in the province. But controversy surrounds the use of FSA results to rank the effectiveness of schools. Critics argue that such rankings reflect the socioeconomic characteristics of the school's community, not the school's performance itself, unfairly giving lower rankings to schools in disadvantaged neighbourhoods.

In an earlier work,[1] I developed a method to measure the influence of socioeconomic factors on educational outcomes in elementary schools in Ontario. That effort showed that the critics are partly right: 40 percent to 50 percent of the variation in schools' assessment scores (averaged over many tests over many years) is associated with variation in the schools' socioeconomic environment. It is reasonable to infer that much of the remaining variation reflects factors at the schools themselves — the principals, the teachers and the other staff. Adjusting test scores to remove the influence of these socioeconomic factors yields measures of relative school performance that are much more representative of a school's actual effectiveness than traditional rankings. This *Commentary* and the associated database apply this analysis to B.C. schools that teach Grade 4 and Grade 7, using the last three available school years of FSA results, 2003/04, 2004/05, and 2005/06.

The analysis I undertake to make a fairer and more useful comparison among British Columbia schools proceeds in three steps. Step One uses the FSA results to create a multiyear measure of achievement that is comparable across schools. Step Two measures the social, economic and educational characteristics of the student body at each school over those three years. Step Three identifies which of these social and economic characteristics have the strongest relationship to FSA test results, and then uses these relationships to predict how well each school would be expected to perform on the FSA achievement measures given the school's socioeconomic characteristics as measured in Step Two. The difference between this predicted achievement score and the actual FSA result provides a measure of the school's relative performance because

I would like to thank the Ministry of Education and Edudata for their help in obtaining the data. I would also like to thank Yvan Guillemette, Bill Robson, Finn Poschmann and the other reviewers for their comments on this paper. All errors are my own.

1 David Johnson, 2005, *Signposts of Success — Interpreting Ontario's Elementary School Test Scores*. Toronto: C.D. Howe Institute.

the variation in achievement scores associated with socioeconomic factors has been removed.

The purpose of such an analysis is not to rank the schools in British Columbia from best to worst, but rather to seek out and create more useful indicators of performance. The results point to schools where teachers, principal, staff and community create an environment that produces high test results relative to other schools with comparable groups of students. Stakeholders may well ask what is done at these schools that could be emulated by others so that all students have an opportunity for better learning.

The rest of this *Commentary* explains the methodology in more detail. The results of the analysis — performance measures for about 1,000 schools that teach Grade 4 and Grade 7 in British Columbia — can be found on the C.D. Howe Institute's website at www.cdhowe.org.

Step One: Measuring School Performance

British Columbia conducts standardized achievement tests through its Foundation Skills Assessment for students in Grades 4 and 7. The focus of this analysis is on students who meet or exceed expectations on the FSA. The remaining students either do not meet curriculum expectations or do not participate in the assessment process. Both groups are important to consider. The FSA tests are designed to be comparable from year to year, and the analysis makes the assumption that for the three academic years studied, 2003/04, 2004/05 and 2005/06, the meaning of meeting or exceeding expectations remains constant.

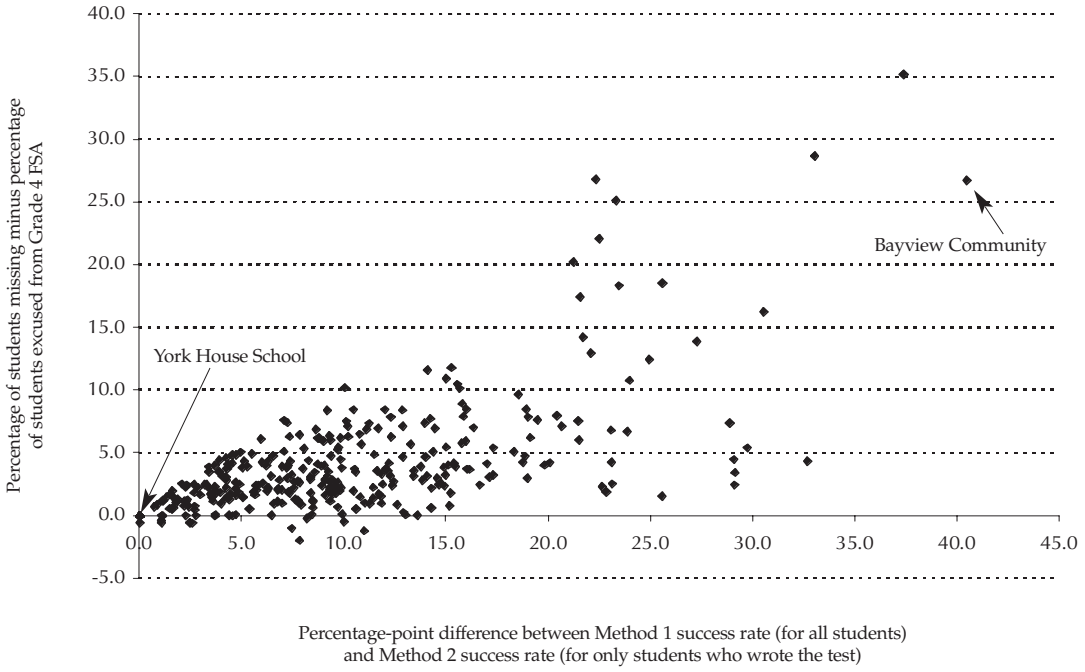
In my methodology, percentages are calculated from individual student data on the three assessments — numeracy, reading and writing — over the three academic years. The total number of assessments (of all three types) in a grade at a school meeting or exceeding expectations is added over the three years and divided by three times the number of students in that grade at the school. This number is then converted into a percentage. However, this percentage is calculated only if at least 45 students attended that school in Grade 4 or 7.[2]

Teachers frequently and correctly point out that the composition of a class varies from year to year. Therefore, using three years of results allows a better measure of what is happening at the school rather than capturing uninteresting random variation in class composition between years. (However, if a school has been in operation for only one or two years, an equivalent count of successful FSA results is made using the results from those years as long as the total number of students is at least 45 students.)

The measure of achievement created above is called here the Method 1 success rate. It includes all students in the denominator in the calculation of the percentage of successful assessments. This makes sense because in a successful school system, all students would meet or exceed expectations at each grade level on all assessments. However, some students do not participate in the FSA process. Some of these students are formally excused for one or more assessments. An equal or far greater number of students are simply missing. In B.C. they are identified in school FSA reports as either

2 Schools operated by district number 93 are not included in the analysis. This district has too few schools with at least 45 students in the relevant grades to allow a meaningful analysis.

Figure 1: *The Role of Missing Students in Greater Vancouver — Grade 4 Results*



Source: Author's calculations.

"not represented" or as participants who did not provide enough responses to be assessed against provincial standards.

Some analysts use a second measure of success in which the number of all assessments that meet or exceed expectations is divided by the total number of assessments either actually written or with sufficient material to grade. I refer to this as the Method 2 success rate. Since the denominator for the percentage calculated in Method 2 is smaller than the denominator for the percentage calculated in Method 1 and the numerator for both expressions is the number of FSA results meeting or exceeding expectations, Method 2 has a higher numerical value if even one student at the school is excused or missing. In very rare cases, it may make sense to actually use the Method 2 success rate. Schools may vary considerably in their proportion of students who, for legitimate reasons, are formally excused from participation in the FSA process.

There are other problems with using Method 2 success rates. There is considerable evidence in both the Grade 4 and Grade 7 FSA data that students who would not meet or exceed expectations are also the students who are systematically missing but not excused from the FSA results. This evidence is presented in Figure 1 for Grade 4 students in the districts around Vancouver (district numbers 38-41 and 43-45). The horizontal axis of Figure 1 measures the difference between the Method 2 and Method 1 success rates. The vertical axis measures the difference between the percentage of students missing and the percentage of students excused. Notice this value is positive

for virtually all schools.[3] This is expected because some students are ill or absent for other legitimate reasons when the FSA is administered.

Only at York House School is both the percentage of missing students and excused students zero. Thus, at York House School, Method 1 and 2 success rates are identical. For all other schools, the Method 2 success rate is higher. The interesting part of Figure 1 is the strong positive relationship between an increase in the gap between the percentage of missing students and the percentage of excused students and the gap between the Method 2 and Method 1 success rates. The conclusion from Figure 1 and similar evidence for Grade 7 is that there is a mechanism in place in the system as a whole and for some schools in particular where the students who are missing and not excused are much more likely to be the students who would not have met or exceeded expectations on the FSA assessments. To put it bluntly, if a school is able to arrange for all students who would not meet or exceed expectations to be missing even when they are not excused, the Method 2 measure of success would be, by definition, 100 percent while the Method 1 measure of success would be much lower. Any methodology that focuses on the grades achieved only on tests written or on the Method 2 success rate is suspect if schools have the ability to influence the composition of the students who are missing. Figure 1 suggests this ability is used by some schools. The problem of missing students also appears to be growing over time.

The percentage of missing students increased in the 2005/06 academic year from the previous two academic years. In both 2003/04 and 2004/05, the percentage of students missing was between less than 10 percent on all assessments in both grades. In 2005/06 the percentage of missing students was over 10 percent on all assessments in all grades. This increase in the percentage of missing students means Method 2 success rates are not comparable across the three years. Average grades on FSA papers actually written would also not be comparable across these three years if the composition of participating students changes.

Figure 1 makes it clear that there are some schools in the Vancouver area where the number of missing students far exceeds the number of excused students. A prime example is Bayview Community School where only 15 percent of students are formally excused, but just 58 percent of potential FSAs in Grade 4 were written over the three-year period studied. Bayview's Method 2 success rate is 96 percent; that is, of the students who actually wrote the Grade 4 FSA, nearly all succeeded. This group would have a high average percentage score on the FSA. But 42 percent of FSA assessments were not written so the Method 1 success rate, calculated as a percentage of all students at Bayview, is only 56 percent. Method 1 success rates on the FSA are a better measure of school performance.

This phenomenon should raise serious concerns with the use of Method 2 success rates and induce parents and administrators to look at those rates in the context of the percentages of students missing and excused. This is particularly important in British Columbia since eight of the nine inputs to the Fraser Institute School Report Cards prior to 2007 are the percentage grades obtained by students who actually write the assessments. The ninth input was 100 minus the Method 2 success rate. As of the May

3 There are nine schools where the percent excused actually exceeds the percent missing by about one percentage point. Here I believe it is likely that an excused student was added by the time of assessment but this student was not in the student count at the school.

2007 Fraser Report Card an additional variable was used to calculate the Fraser overall rating. That variable (with only a 10 percent weight) measures the percentage of tests not written beyond students formally excused. This is a small step towards using a Method 1 success rate. However 90 percent of the weights in the 2007 Fraser Report Card ratings use only information from assessments written. Because the Fraser Report Card makes such heavy implicit and explicit use of Method 2 success rates, an effective strategy for a school in British Columbia that wishes to improve its Fraser Institute ranking is to ensure that low-performing students are missing or write exams that do not contain meaningful material on assessment day. The data in Figure 1 strongly suggest this is indeed a practice at some schools.

There are three clear ways to substantially improve the analysis and use of FSA data in British Columbia, ways that are implemented in the rest of this *Commentary*. The first contribution of the study is to use multiple years of data and thus appropriately reduce emphasis on year-to-year results. The second contribution is to place much more emphasis on Method 1 results so that all students are considered part of the FSA process for the purpose of comparing schools. The third contribution is to take into account the social, economic and educational context of the school in comparing schools using Method 1 success rates. The next two sections of the paper address this critical third task.

Step Two: Creating a Socioeconomic Picture of a School

How can the socioeconomic characteristics of students at a school be measured? There are two sources of data. One source links student home locations to data in the 2001 census. The details of this process can be found in Chapter Three of *Signposts of Success* (Johnson, 2005) and are summarized here. Everyone lives in a designated Dissemination Area (DA), the smallest geographic unit of the census with an average of 584 persons. There are about 7,000 DA units in British Columbia. The characteristics of the average person or family unit that lives in each small geographic area are part of the information reported in the census. It is possible to calculate how many students at a particular school in each year lived in each Dissemination Area. With this information, the process of constructing a census-based school profile is straightforward.

For example, say a given school has 10 students who live in two different DAs. Six students live in DA 1 and four in DA 2. In DA 1, 80 percent of the families are single-parent families, and in DA 2 only 20 percent are single-parent. The percentage of lone parents at this school is calculated as (0.6 x .80) + (0.4 x .20) = 56 percent. The larger the school, the more accurately the school profile will resemble the profile of students who wrote the achievement tests over the three years studied.

Other census-based variables that describe each school population are: the percentage of persons who moved in the past calendar year; the percentage of persons living in detached dwellings; the percentage who immigrated to Canada in the past five years; the percentage who speak an official language as their mother tongue; the percentage who are aboriginal; the unemployment rate of adults with children; two measures of income (average household income and average family income); and, finally, three measures of education (the percentage of those over 20 years of age

without a high-school diploma, the percentage over 20 years old who have some university education, and the percentage over 20 with a university degree).

There is a second source describing the social, economic and educational characteristics of students participating in the assessment process in British Columbia. This B.C. data is unique in the Canadian context and makes the process of setting school results in comparable environments more precise than in other provinces.[4]

In B.C., each report of a student outcome includes identifiers that record some individual characteristics. A student is male or female; an English-as-a-Second-Language student; or an aboriginal. He or she speaks English at home; is in French immersion; is a designated special education student; and, finally, may be designated as gifted. If an individual student is formally excused from each FSA, then this information is also recorded. From this data it is straightforward to calculate the percentage of students in each grade at each FSA school that are male, aboriginal, enrolled in ESL, French Immersion or in various special education categories. These percentages directly describe the students eligible for the FSA in the three years studied. The school profiles derived from the FSA and census data are on the C.D. Howe website at www.cdhowe.org and are available for use by other interested researchers and school boards.

The final step in this *Commentary's* school evaluation process asks how the measures of the social, economic and education composition of the class relate to the school's Method 1 success rate. Only then can the performance of similar schools be compared.

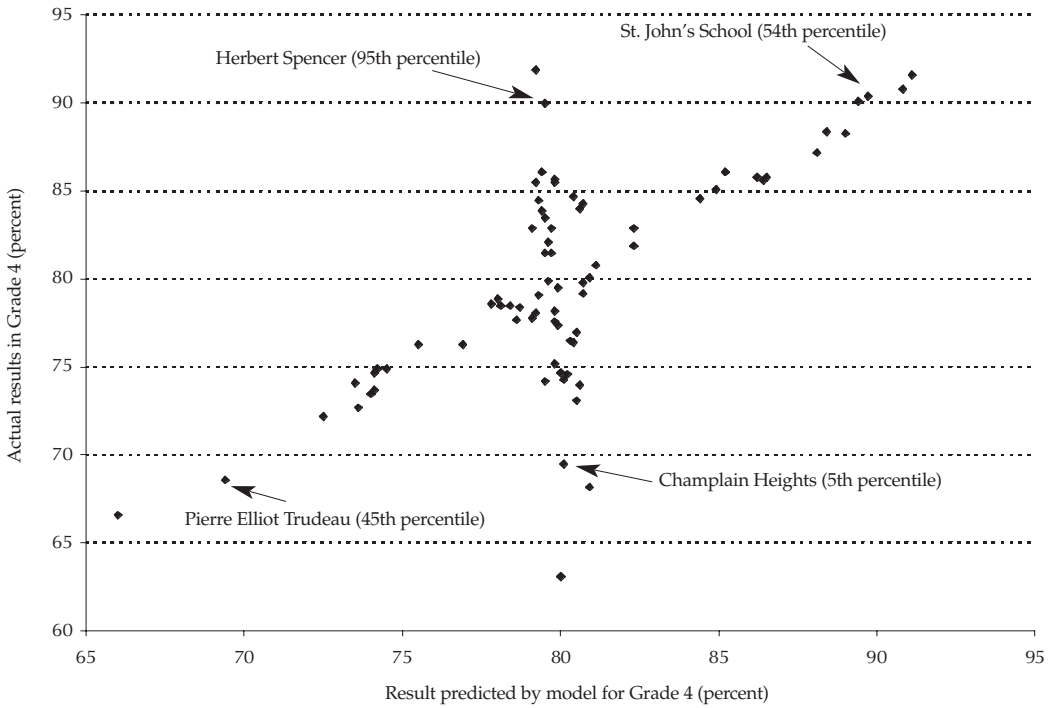
Step Three: Comparing School Performance

More detail on the methodology used in the analysis can be found in Chapters 4 and 6 of Johnson (2005). The methodology is summarized here with the help of Figure 2. On the vertical axis is the percentage of all students in Grade 4 meeting or exceeding expectations over the three-year period; that is, the Method 1 success rate discussed above. On the horizontal axis are values of an index of social, economic and educational indicators (the details of their calculations are found in the Appendix). The dots on the diagram are 77 of the 294 schools in the Vancouver area. (There are 352 schools reporting data from these districts, but many of the 58 excluded schools have less than 45 students in one or more of the two grades. In six excluded schools, more than 50 percent of the students are missing, making it difficult to draw conclusions from the data.

Of the 77 included schools, some of these 77 schools are on the upward sloping line representing the statistical relationship between the social, economic and educational variables and the percentage of all students at the school that meet or exceed expectations (the Method 1 success rate). In other words, the values along the upward sloping line establish a "predicted" success rate for a given school and grade, given the environment in which that school operates. That predicted pass rate allows a

4 The B.C. data allows analysis at the level of the individual student and enables educators to understand better the links between individual and school results. Since the focus of this study is on understanding what the FSA results can and cannot say about the quality of a school, making further use of the individual data is left to future research.

Figure 2: *Grade 4 Results for some Greater Vancouver Schools — percent of all students meeting or exceeding expectations*



Source: Author's calculations.

comparison of the success rate actually achieved at each school with the rate at schools that operate in similar environments.

If a school is on that upward-sloping line, then this school's performance is the same as other schools in the province with a similar socioeconomic mix of students. At one such school, Pierre Elliott Trudeau, only 68 percent of its Grade 4 students met or exceeded expectations over the three years. This is well below the provincial average result and below the average of results illustrated in Figure 2. But does this make P.E.T. a poor school? According to our relative performance measure, the answer is no. Based on its mix of students, P.E.T. is expected to have 69 percent of its students meet or exceed expectations. Thus, the relative performance of students at P.E.T. is the same as other schools in similar environments.

If one travels up the upward sloping line past its intersection with the vertical line, one arrives at St. John's School where slightly more than 90 percent of all students meet or exceed expectations. Does this mean St. John's is a better school than P.E.T.? The answer again is no. At St. John's, the characteristics of the student body suggest that 90 percent of its students will meet or exceed grade level, and this is exactly what happened. St. John's and P.E.T. are both average schools in the sense that the percentage of their students who meet or exceed expectations is exactly what would be predicted by the composition of their student body.

When is a school strong or weak in this analysis? For the answer, consider the 40 or so schools along the vertical line (some schools are both on the upward sloping and

vertical lines). All of these schools have a mix of students whose socioeconomic characteristics predict that between 79 and 81 percent of all students would meet or achieve grade-level expectations.[5] The actual percentage of students at these schools who meet or exceed provincial expectations varies widely, from 63 to 91 percent. It is thus fair to compare the results of the schools along the vertical line as indicators of how good a job their staff is doing. For example, it is reasonable to conclude that the Herbert Spencer school, where 90 percent of all students over three years met or exceeded expectations, is a better school than Champlain Heights where, with a similar mix of students, only 70 percent of all students attained the same standard. A reasonable question to ask — one that should be the objective of future research — is what is happening at the Spencer school that allows its students to do better, and what does or does not happen at the Champlain school to explain its poorer relative performance?

A Numerical Presentation of Relative School Performance Measures

All schools in British Columbia where Grades 4 and 7 are taught could be placed on a diagram like Figure 1. With two grades and 1,000 schools, however, the mass of dots would be unreadable. Instead, the data are presented in a large table on the C.D. Howe Institute's website. Presenting these kinds of numbers is necessary to make fair comparisons among schools and to understand how the different measures of success are and are not relevant. The first three columns identify the school by district number, name and Ministry of Education school number. The next six columns relate to Grade 4 FSA results (the following 6 columns relate to Grade 7 results). The columns of data begin with the total enrolment at the school in Grade 4 over the three academic years, i.e. the total number of students potentially eligible to write the Grade 4 FSA. The next column is the Method 1 success rate. The third column is the predicted Method 1 success rate for that school, the values represented by the upward sloping line in Figure 1.

With this information, it is straightforward to calculate the differences between the predicted and actual success rates for every school. If the actual rate is higher than the predicted rate, this school is above average. The table does not present this calculation. Instead, it presents an even better measure of a school's quality than the simple difference between the school's actual score and predicted score where a positive value says the school is above average and a negative value says the school is below average. The amount a school is above or below its predicted score is transformed and expressed as a percentile, a rating from 0 to 100. A percentile score near 50 indicates that, compared to schools with students that have similar social and economic

5 The choice to look at schools where the predicted score lay between 79 and 81 percent is not important. I could have looked at schools whose predicted scores lay between 74 and 76 percent and the diagram would convey the same information. If I had chosen to look at schools where predicted scores ranged between 74 and 76 percent, the vertical line would simply be to the left of the current vertical line. There is a vertical line at each value of the predicted score when you look at all 1,000 schools with Grade 4 in British Columbia or at all 294 included schools in the Vancouver area.

characteristics, a school is average: half the schools are better and half are worse. On the other hand, a percentile score of 90 says that a school is better than 90 percent of schools whose students have similar socioeconomic characteristics. This would be a good, indeed a great, school. The percentile is a rating of a school using Method 1 and comparing results at that school to results at schools where the social, economic and educational characteristics of the students predict similar FSA results. A large difference in percentile ratings between schools means that it is reasonable to assume there is a large difference in school performance and quality between schools with similar mixes of students. This is a fair and interesting comparison of schools. There is a little more to be said about the important differences between Method 1 and Method 2 success rates by looking at the last three columns of data pertaining to Grade 4.

These columns begin with a presentation of the Method 1 success rate. But a discerning reader may notice that a successful school in terms of a high Method 1 or Method 2 success rate often draws its students from pools that are more likely to succeed. The predicted Method 1 success rate may be high for such a school but the percentile rating of the school will be 50 or even less, and the school will be revealed as average or simply comparable to other schools that draw their students from similar populations. However the last two columns under the Grade 4 heading present a second reason why a school might have a low Method 1 success rate and a high Method 2 success rate. One column shows the percentage of students formally excused from the annual testing. This value varies from zero to 40 percent. A school where a large percentage of students are formally excused will have a relatively low Method 1 success rate. In this case, and only in this case, assessment of the school success rate might well focus usefully on Method 2 success rates and percentiles. There are only a few schools in this category. In 176 schools, more than 10 percent of the Grade 4 students were formally excused from the tests; in only 20 schools, more than 20 percent of the students were excused. The last column pertaining to Grade 4 shows the percentage of students missing (including the percentage excused). You can identify schools where the percentage of students missing is much larger than the percentage of students formally excused. In these schools the Method 2 success rate is much higher than the Method 1 success rate, and the evidence presented in Figure 1 strongly suggests that students who are missing but not excused are students who would not have met expectations on the FSA. For these schools in particular, and for all schools in general, Method 1 success rates and particularly the percentiles provide a better measure of the success of a school than Method 2. The role of missing but not excused students is an important and thus far neglected part of the analysis of FSA results in British Columbia. The table then displays similar information for Grade 7 FSA results in the last six columns of data.

Percentile ratings for about 1,000 British Columbia schools in one of the two grades are found in the Table accessed on the C.D. Howe Institute website. Returning to Figure 2 will help the reader better interpret the percentile ratings. In the four schools labeled by name in Figure 2, percentile scores are shown in parentheses. P.E.T. and St. John's have a percentile score near 50, meaning that they produce the same percentage of students meeting or exceeding expectations as do other schools that operate in a similar socioeconomic context. Herbert Spencer scores at the 95th percentile, while Champlain registers at the 5th percentile. Herbert Spencer produces results much

better than those at Champlain although the socioeconomic mix at both schools predicts the same result.

Comparing schools ideally becomes an exercise in the interpretation of percentile numbers. If one school has a rating in the 95th percentile and another school has a rating in the 5th percentile, it is very likely that the staff at the 95th percentile school is doing a better job than the staff at the 5th percentile school. Differences of this magnitude are worth investigating. They occur among both Grade 4 and Grade 7 results.

There are many blank entries in the table on the web for various reasons. First, some districts have elementary schools that end at Grade 5 or Grade 6. Second, schools where less than 45 students over three years wrote a Grade 4 or Grade 7 FSA test are not included in the comparisons. It is simply not valid to rate schools based on small numbers of student results. Finally, there are a few schools where, although there are the requisite 45 students, more than 50 percent of the students are missing. These schools are not included.

Conclusion

The foregoing analysis solves, to a very large degree, the problem of using FSA results to make a fair comparison of schools. It uses a number of years of data to remove year-to-year fluctuations. It uses a large sample of students. It compares schools where students are from similar socioeconomic backgrounds. It considers the problem of missing students and the differences between including all students in the comparisons and using only students who actually write the examinations. There remain substantial differences in the performance of students in FSA tests across similar schools. There are excellent schools and they stand out. The next task will be to discover what happens at these excellent schools that produces excellent results, and how best practices could be transferred to other schools.

Appendix: The Creation of the Index of Social and Economic Indicators

The values for the Index of Social and Economic Factors plotted in Figure 1 and presented in the data table are the fitted values from a regression. The dependent variable in the regression is the Method 1 success rate. Table A1 presents the coefficients on the independent or right-hand-side variables in these regressions.

All the social, economic and educational variables from the school data and the census data were considered as potential explanatory variables for differences in FSA results. From the list of student-based variables, the five at the top of Table A1 were significantly related to FSA outcomes in either Grade 4 or Grade 7. Among the census-based variables, the three that appear at the bottom of Table A1 are those most strongly associated with school achievement results in the context of the school data.

The same variables are used to predict Grade 4 and Grade 7 FSA results. This both simplifies the presentation and highlights important differences in the relationships between specific variables and results across the two grades.

The coefficient values in Table A1 have a clear interpretation. The coefficient value negative 0.051 (standard error 0.034) in the column labeled "Coefficient in Grade 4 Regression (standard error)" means that, everything else being equal, if school A has one percentage point more boys than school B, the percentage of students who meet or achieve expectations is predicted to be 0.051 percentage points lower at school A than school B. The standard error value in parentheses, 0.034, is large relative to -0.051. This means that the value -0.051 is not precisely estimated. In fact, the partial relationship between gender and FSA results in Grade 4 just described is not statistically significant at conventional levels. But the gender variable remains in the model because the relationship between results and gender is much stronger and is statistically significant in the analysis of Grade 7 FSA results. This coefficient is found in the last column of the first row of coefficients. In the analysis of Grade 7 FSA results, if a school-grade combination has one percentage point more boys, then it is predicted to have 0.131 fewer percentage points of all students meeting or exceeding expectations. Notice the standard error of this coefficient is 0.043, and that this is a small value relative to 0.131. This result is statistically significant at conventional levels. The remaining coefficients in the relationships are interpreted in the same way.

The effect of an additional one percentage point of ESL students has a negative effect on the Method 1 success rate in Grade 4, but has no significant relationship to FSA success rates in Grade 7. The difference in the ESL relationship to results by grade level is important. It implies, other factors being equal, that by Grade 7 an ESL-designated student has a similar or better success rate than a non-ESL student. This may reflect considerable success in the education system in handling ESL students, many of whom are immigrants.

The next row of coefficients measures the influence of an additional percentage of aboriginal students. Here it is clear that having a larger proportion of aboriginal students is predicted to reduce FSA success rate in both grades. The impact is profound. In Grade 4, an additional one percentage point of aboriginal students predicts a 0.325 percentage point reduction in Method 1 success rates; in Grade 7 the predicted reduction is 0.436 percentage points. The coefficients describing the relationship between the aboriginal component of the school and FSA results are

Table A1: *The Regressions Behind the Index of Social and Economic Factors*

Variable	Coefficient in Grade 4 Regression (standard error)	Coefficient in Grade 7 Regression (standard error)
Percentage of boys	-0.051 (0.034)	-0.131 (0.043)
Percentage of ESL students	-0.106 (0.022)	-0.027 (0.070)
Percentage of aboriginals	-0.325 (0.035)	-0.436 (0.038)
Percentage of English-speaking students	0.048 (0.017)	0.083 (0.020)
Percentage of special-education students	-0.486 (0.067)	-0.440 (0.070)
Variables above derived from FSA individual student data. Variables below derived from student home location and census data		
Percentage who moved in the past year	-0.098 (0.045)	-0.082 (0.054)
Percentage of lone parents	-0.111 (0.027)	-0.070 (0.032)
Percentage over 20 years old who have some university	0.162 (0.022)	0.229 (0.030)
Number of schools in the regression	1027	828
Explanatory power of the regression	0.55	0.60

Source: Author's calculations; 2001 Census; Edudata Canada; British Columbia, Ministry of Education Foundation Skills Assessment 2003/4, 2004/5, 2005/6.

clearly different in Grade 4 and Grade 7. Thus the composition of the school is again predicted to affect FSA results differently in different grades.

The same conclusion can be drawn about the relationship between FSA results and the percentage of students who speak English at home. In Grade 7, the predicted increase in FSA results as a result of a one-percentage-point increase in English-speaking students is 0.083 percentage points; that value is 0.048 percentage points in Grade 4. Speaking English at home matters slightly more at the higher grade level.

Finally, the last variable derived directly from the individual student level that enters the prediction equation is the percentage of students who are designated special education students. Here, the Grade 4 and Grade 7 equations yield very similar coefficients. A school that has one percentage point more students designated special education students than another reduces the predicted success rate in Grade 4 and in Grade 7 by about one-half of a percentage point. It seems reasonable that when more students are designated as needing special help, FSA results are expected to be lower. However, it is discomforting to have such a variable that schools can clearly influence play a large role in establishing the predicted FSA result. This comment should recall to mind the discussion around the evidence that schools almost certainly influence who is missing for FSA tests in order to improve results reported based on Method 2. However, the percentile ratings for schools are virtually identical when the percentage

of special education students is removed from the prediction equation. Indeed, a wide variety of combinations of variables in the prediction equations yield very similar percentile ratings.* For example, it is possible to generate good census variable proxies for the percentage of ESL (or alternatively, English as a home language) students as well as for aboriginal students. Such a census-based variable could be used in the prediction equation instead of the ESL variable based on the grade-school data.

The three census-based variables listed in Table A1 and not available in the student-based data are clearly related to FSA success rates. If a school community has one percentage point more people who have moved in the last year, success rates are predicted to fall by 0.098 percentage points in Grade 4 and 0.082 percentage points in Grade 7. If one percentage point more children come from lone-parent families, success rates are predicted to be lower by 0.111 percentage points in Grade 4 and by 0.070 percentage points in Grade 7. Finally, there is a strong relationship between a measure of parental education and FSA success rates. The proxy variable for parental education is the most useful census variable added to the model. If one percentage point more of adults over 20 in the school community has some exposure to university, success rates are predicted to rise by 0.162 percentage points in Grade 4 and by 0.229 percentage points in Grade 7. These are large and precisely estimated increases. The association of strong school results and the measure of parental education increases between Grade 4 and Grade 7. This is similar to findings in both Alberta and Ontario.

The last two rows in Table A1 are marked "number of schools in the regression" and "explanatory power of the regression." The former is the total number of schools with at least 45 students over the three years in the grade studied. Schools where more than 50 percent of students were missing were also excluded.

The explanatory power of the regression number, which must fall between 0 and 1, is an indicator of the strength of the overall relationship between socioeconomic factors and school achievement results. If this value were 1.0, then the dots representing all schools would fall along the upward sloping line in Figure 2. There would be no schools along the vertical line. All of the variation in the percentage of students meeting or exceeding expectations at various schools would be associated with the students' socioeconomic and educational background.

If the "explanatory power of the regression" was, instead, zero, then variation in the social and economic background of students would not be associated with variation in achievement test results. That would surprise everyone. The two regression values in Table A1, 0.55 for Grade 4 and 0.60 for Grade 7, indicate that 55 percent and 60 percent of the variation in the percentage of Grade 4 and Grade 7 students that meet or exceed expectations is associated with the set of social, economic and educational factors in the models estimated in the table. The remaining variation is associated with other factors that are not observed.

It is reasonable to infer that much of the remaining variation reflects factors at the schools themselves — the principals, the teachers and the other staff. In other words, it

* The author would be happy to provide an equivalent percentile rating using only census variables in the predication equation on request. Dropping the special education variable leads to virtually identical prediction results. It is likely that the designation of a student as needing special education help occurs for a wide variety of reasons that are related to the social and economic background of the students that are well-measured in the census-based variables.

reflects the quality of the teaching and teaching environment. The school performance measures are based on this residual. It accounts for 40 to 45 percent of the variation between school results. There is, therefore, substantial evidence that how schools do their job really matters for student outcomes. Excellence can then be identified and celebrated.

References

Johnson, David. 2005. *Signposts of Success: Interpreting Ontario's Elementary School Test Scores.* Policy Study 40. Toronto: C.D. Howe Institute.

NOTES

NOTES

Recent C.D. Howe Institute Publications

January 2008	Laidler, David. "Fixing the Exchange Rate: Still a Bad Idea." C.D. Howe Institute e-brief.
January 2008	Robson, William B.P. "Safe Harbours: Providing Protection for Canada's Money-Purchase Plan Sponsors." C.D. Howe Institute Backgrounder 110.
December 2007	Robson, William B.P. "Time and Money: The Challenge of Demographic Change and Government Finances in Canada." C.D. Howe Institute Backgrounder 109.
December 2007	Chen, Duanjie. "Flaherty's Missed Opportunity." C.D. Howe Institute e-brief.
December 2007	Guillemette, Yvan. "Chronic Rigidity: The East's Labour Market Problem and How to Fix It." C.D. Howe Institute e-brief.
December 2007	Kotowski, Maciej. *Insuring Canada's Exports: The Case for Reform at Export Development Canada*. C.D. Howe Institute Commentary 257.
December 2007	Koning, John-Paul. "Why the Bank of Canada Act Needs Updating: A Lesson from the Sub-Prime Crisis." C.D. Howe Institute e-brief.
November 2007	Jaccard, Mark. *Designing Canada's Low-Carb Diet: Options for Effective Climate Policy*. Benefactors Lecture 2007.
November 2007	Banerjee, Robin. "Deals on Wheels: An Analysis of the New Federal Auto Feebate." C.D. Howe Institute Backgrounder 108.
November 2007	Freedman, Charles and Clyde Goodlet. *Financial Stability: What It Is and Why It Matters.* C.D. Howe Institute Commentary 256.
November 2007	Guillemette, Yvan and William B.P. Robson. "Realistic Expectations: Demographics and the Pursuit of Prosperity in Saskatchewan." C.D. Howe Institute Backgrounder 107.
October 2007	Macmillan, Kathleen E. and Patrick Grady. "A New Prescription: Can the BC-Alberta TILMA Resuscitate Internal Trade in Canada?" C.D. Howe Institute Backgrounder 106.
October 2007	Richards, John. *Reducing Poverty: What has Worked, and What Should Come Next*. C.D. Howe Institute Commentary 255.
October 2007	Guillemette, Yvan. "Breaking Down Monopolies: Expanding Choice and Competition in Education." C.D. Howe Institute Backgrounder 105.
September 2007	Mintz, Jack M. *The 2007 Tax Competitiveness Report: A Call for Comprehensive Tax Reform.* C.D. Howe Institute Commentary 254.
September 2007	Chen, Duanjie, Jack Mintz and Andrey Tarasov. "The High-Tax Handicap: How the World's Major Economies Shoot Themselves in the Foot with High Corporate Income Tax Rates." C.D. Howe Institute e-brief.
September 2007	Johnson, David. "School Grades: Identifying Alberta's Best Public Schools." C.D. Howe Institute Backgrounder 104.
August 2007	Vaillancourt, Francois, Dominique Lemay and Luc Vaillancourt. "Laggards No More: The Changed Socioeconomic Status of Francophones in Quebec." C.D. Howe Institute Backgrounder 103.
août 2007	Vaillancourt, François, Dominique Lemay and Luc Vaillancourt. "Le français plus payant : L'évolution du statut socio-économique des francophones au Québec." Bulletin de recherche 103.
August 2007	Robson, William B.P. "Stuck on a Spoke: Proliferating Bilateral Trade Deals are a Dangerous Game for Canada." C.D. Howe Institute e-brief.
July 2007	Smart, Michael. *Lessons in Harmony: What Experience in the Atlantic Provinces Shows About the Benefits of a Harmonized Sales Tax.* C.D. Howe Institute Commentary 253.
July 2007	Laidler, David. *Better Late Than Never: Towards a Systematic Review of Canada's Monetary Policy Regime*. C.D. Howe Institute Commentary 252.
July 2007	Chen, Duanjie, Jack Mintz and Andrey Tarasov. "Federal and Provincial Tax Reforms: Let's Get Back on Track." C.D. Howe Institute Backgrounder 102.
June 2007	Koeppl, Thorsten V., and James MacGee. *Branching Out: The Urgent Need to Transform Canada's Financial Landscape and How to Do It.* C.D. Howe Institute Commentary 251.
June 2007	Jaccard, Mark, and Nic Rivers. "Estimating the Effect of the Canadian Government's 2006-2007 Greenhouse Gas Policies." C.D. Howe Institute Working Paper.
June 2007	Jaccard, Mark, and Nic Rivers. "Estimating the Effect of the Canadian Government's 2006-2007 Greenhouse Gas Policies." C.D. Howe Institute e-brief.

C.D. Howe Institute
67 Yonge Street Suite 300
Toronto, Ontario
M5E 1J8