The Alberta Journal of Educational Research *Vol. 56, No. 1, Spring 2010, 95-98*

Research Notes

and

Robert E. Krider
Simon Fraser University

Assessing the Socioeconomic Correction in the Fraser Institute Report Cards

Introduction

The ranking of schools in Alberta, British Columbia, Ontario, and Quebec by the Fraser Institute has been publicly criticized on a variety of dimensions. In response to one dominant criticism, that the rankings are heavily influenced by factors other than school quality, the Fraser Institute now includes an attempt to control for socioeconomic factors (Cowley & Easton, 2008). The result, the socioeconomic indicator (SE), is offered as a measure of whether the school is doing better or worse than expected given the socioeconomic characteristics of the school population. Ideally, this should provide parents and educators with a more nuanced assessment of intrinsic qualities of schools. In the research reported here, we assess the effectiveness of this control by exploring measures that may be related to the reported better-or-worse indicator. We find that the control overcompensates on some dimensions, for example, household income, and undercompensates on others, for example, the percentage of single-parent families. This provides additional context in interpreting the measure and may assist in the development of better measures.

Data Analysis

The school academic performance and descriptive variables were taken from the *Report Card on Secondary Schools in British Columbia and Yukon* published by the Fraser Institute for the 2007-2008 school year. The description of how these variables were derived can be found in the report. The socioeconomic variables were taken from the Statistics Canada *Census 2006 Community Profiles*, which are freely available online. The data are 20% sample data at the district municipality level. These variables were chosen for their potential effect on a student's home life.

In the Fraser Institute's report, the academic performance variables are used in a model that outputs an overall rating between 0 and 10 for each school. The Institute's SE indicator is based on a regression between the overall rating and the mean number of years of education of parents in the school's neighborhood

Dustin Dunsmuir is a graduate student in the School of Interactive Arts and Technology. His research involves techniques for analyzing large amounts of data through interactive visualizations. He can be reached at dtd@sfu.ca.
Robert Krider is a professor of marketing in the Faculty of Business Administration. His research involves understanding the mechanisms of competition and customer choices in a range of contexts from movies to schools. He can be reached at rkrider@sfu.ca.

taken across all schools in the province. This relation is used to predict an expected rating for the school based only on parents' education. From these two ratings the report gives a difference (which we call the SE indicator) between the actual rating and the expected SE rating, which is positive if the school did better than expected and negative if the school did worse, given parents' education. We implement this better/worse output as a binary dependent variable and model it as a function of the variables in Table 1.

The Fraser Institute report contains 298 schools. Unfortunately, some schools were not used in our analysis due to missing data. Schools that did not list an expected SE rating (which included all Yukon schools) were necessarily removed from consideration. This left 280 schools that were used in modeling. A small percentage of these cases were missing the values for the delayed advancement rate, English gender gap, or math gender gap, and these were set to the mean value derived from all the other cases so as not to lose the data on the other variables for those schools. The only strong correlation found among the predictor variables was a Pearson's correlation of –0.91 between the mean exam mark and the exam failure rate.

Our exploratory approach used three types of models. The first was a recursive partitioning tree model used to understand which variables had the greatest effect on the target variable. The second was a neural network. The neural network is too complex to interpret on its own, but is valuable for determining the potential predictive power and possible nonlinear and interaction effects of the predictors. Nonlinear relationships were found, and these relationships prompted the calculation and inclusion of 12 transformed (square or log) predictor variables. Insights from the tree and neural network models informed the final two models, logistic regressions, estimated stepwise using Akaike's Information Criteria (AIC) to select the meaningful set of predictors. The first of these uses only predictors that are not used to construct the overall Fraser Institute score. Table 2 shows the results of this analysis. The second uses all variables in Table 1 including those contributing to the overall score, with results shown in Table 3.

Table 1
List of Possible Predictor Variables Explored in the Analysis Grouped by Type

Academic Performance	*School Descriptive*	*Socioeconomic*
School mark vs. exam mark difference	Grade 12 enrollment	Percent of families with lone parent
Mean exam mark	Public or independent	Mean number of people in a family
Exam failure rate		Median family income
English gender gap		
Math gender gap		
Graduation rate		
Delayed advancement rate		

Table 2
Unstandardized and Standardized Logistic Regression Coefficients for Effects of Socioeconomic Variables on the Binary Socioeconomic Indicator

	b	*ß*
Percent of lone parent families	0.787	3.00*
(Percent of lone parent families)^2	–0.030	–3.44***
(Median family income)^2	–0.030	–2.83**
School type: Public	–3.28	–2.49*
Intercept	–18.7	–4.82**
McFadden R^2	0.18	
N	280	

*p<0.05; **p<0.01; ***p<0.001.

Results

In Table 2, the first two coefficients describe a surprising inverted *U* effect of the lone-parent families. Simulation helps to describe the details. As the percentage of such families increases to 13%, the likelihood of the SE indicator being positive increases to a maximum and then starts to decline; at 27%, the effect of lone-parent families is neutral. Increasing median family income causes the SE indicator to decrease nonlinearly, implying that the Fraser Institutes' compensation is actually overcorrecting for income and at an increasing rate. Private schools are more likely to be above expectations than public schools.

Because the overall rating is an ad hoc weighting of input measures, we ran a second analysis that includes those variables as well. Any additional significant predictors found might be useful in constructing better weightings, and as covariates may allow the previous predictors to have stronger effects. Note that the resulting stepwise model contained both the average exam mark and the exam failure rate, which are correlated, and thus the average exam mark was removed from this model. This model is shown in Table 3.

Significant predictors of the SE indicator that were used in the Fraser Institute's overall score are interesting and point to opportunities for improving the weighting scheme in the score. Most notably, positive scores on the SE indicator are strongly related to higher graduation rates and to lower absolute differences between exam and school grades. The effect of percentage of lone-parent families in the neighborhood is now consistently negative and strong throughout the range, implying insufficient compensation for this effect, although the income effect is consistent with the previous model. The mean number of family members has a complex but intuitively appealing effect. The marginally significant negative linear relation plus positive logarithmic terms together indicate an inverted *U* with maximum at three members and a negative effect as the family size increases beyond this. Note that a model with a single linear term for family size is unable to capture this pattern. Private schools remain more likely to score better than expected although the strength of the effect is much weaker than in the simpler model.

Table 3
Unstandardized and Standardized Logistic Regression Coefficients for Effects of Student Performance and Socioeconomic Variables on the Binary SE Indicator

	b	*ß*
Delayed advancement rate	–0.0357	–2.13*
Exam failure rate	–0.0792	–1.72
Graduation rate	0.209	3.93***
Percent of lone-parent families	–0.148	–2.27*
Math gender gap	0.120	1.75
(Median family income)^2	–0.0617	–4.47***
Mean number of family members	–53.0	–1.93
Log (Mean number of family members)	154	1.99*
(School vs. exam mark difference)^2	–0.0342	–3.22**
School type: Public	–2.65	–2.49*
Intercept	–18.7	–2.71**
McFadden R^2	0.433	
N	280	

*p<0.05; **p<0.01; ***p<0.001.

Discussion

The Fraser Institute's evaluation of school performance is a difficult and politically charged exercise. The authors of the reports are diligent in providing caveats for the interpretation and use of the report cards and have responded to some criticisms by including a compensation for socioeconomic conditions. The research reported here focuses on this compensation and finds it a useful improvement. We also find some remaining issues, notably the effect of single-parent families, which can be used as a context for a more nuanced interpretation of the results, as well as a direction for improving the measure.

Reference

Cowley, P., & Easton, S. (2008). *Report card on secondary schools in British Columbia and Yukon: 2008 edition*. Retrieved May 26 2009, from: http://www.fraserinstitute.org/researchandpublications/publications/5563.aspx