The Effect of High School Resources on Investment in Post-Secondary Education in Canada

Kathleen Day

MESA2009-2

MESA MEASURING THE EFFECTIVENESS OF STUDENT AID

Canadian Education Project | Queen's University School of Policy Studies | Canada Millennium Scholarship Foundation
Educational Policy Institute | Higher Education Strategy Associates

Toronto, Ontario, Canada - June 2009

www.mesa-project.org

The MESA Project

The Measuring the Effectiveness of Student Aid Project, or the MESA Project, is a four year research effort being conducted by the Canadian Education Project and the School for Policy Studies at Queen's University on behalf of the Canada Millennium Scholarship Foundation. It has been designed to answer the following four questions:

- After graduating from high school, teenagers coming from low-income backgrounds face a choice as to attend college or university, or not. For those who did attend, how do they compare to those who did not?
- Does providing more funding in a student's first few years of further education attract more low-income students to post-secondary education?
- Does providing more funding in a student's first few years of further education make it more likely for low-income students to stay in and graduate?
- Are low-income students different across Canada?

This paper is part of a series of research papers solicited from some of the leading Canadian researchers in the field of post-secondary education; the researchers were asked to write about issues of access and persistence in post-secondary education in Canada. The requirements for the papers were that the researchers use one of several currently-existing Statistics Canada databases or another source of Canadian data. Each of the papers commissioned during this project is available for downloading from the MESA Project website at www.mesa-project.org.

The findings and conclusions expressed in this paper are those of the authors and do not necessarily represent those of the MESA Project or its partners.

The Partners

The **Education Policy Institute** is an international, non-profit think tank dedicated to the study of educational opportunity. Our mission is to provide high-level research and analysis to support policymakers and practitioners and expand educational opportunity to all students. EPI handles overall project management and co-ordination, data privacy & cleaning, and integration of the final results at the end of the project. www.educationalpolicy.org

The **Canadian Education Project** provides research and evaluation expertise in experimental, quantitative, qualitative and mixed methods research approaches. The company has experience working with a broad range of stakeholders including governments (at the federal and provincial levels), secondary and post-secondary educational institutions, elementary and secondary school boards, student groups, non-profit and non-governmental organizations and other stakeholders in the education and public policy arena in Canada and internationally. While much of our work to date deals with students and youth at the post-secondary level, we are increasingly engaging in research at the elementary and secondary levels as well as looking at student mobility through lifelong learning and transitions between K-12 and post-secondary education. www.canedproject.ca

The **School of Policy Studies at Queen's University** (*www.queensu.ca/sps*) is a leading centre for advanced education, research, debate and interaction with the non-academic world in the fields of public administration and industrial relations. Continuing the long-standing commitment of Queen's University to excellence in these areas, they are training the next generation of leaders for life in a global age. Their master's programs link theory with practice to provide students with fundamental knowledge of the economic, political, social and technological changes that are transforming the way we live and the way we work. Students enhance their communication and research skills, and gain new skills in management, policy analysis, economics and quantitative methods. Their graduates are well prepared to contribute to policy-making, human resource management and industrial relations in a variety of public, private and nonprofit organizations. The School for Policy Studies manages the Research Review Committee for the MESA Project, which is responsible for funding contributory research projects that highlight important policy areas of interest.

The **Canada Millennium Scholarship Foundation** is a private, independent organization created by an act of Parliament in 1998. It encourages Canadian students to strive for excellence and pursue their post-secondary studies. The Foundation distributes $325 million in the form of bursaries and scholarships each year throughout Canada. Its objectives are to improve access to post-secondary education for all Canadians, especially those facing economic or social barriers; to encourage a high level of student achievement and engagement in Canadian society; and to build a national alliance of organizations and individuals around a shared post-secondary agenda. The Foundation is funding the MESA Project overall, and has negotiated access to its student administrative lists with each of the provinces on the project's behalf.

www.millenniumscholarships.ca

Abstract

An important question relating to post-secondary education is whether spending more money at the elementary and high school levels to better prepare students for post-secondary education constitutes a more cost-effective means of encouraging investment in post-secondary education than providing more money to students and/or institutions at the post-secondary level. This study uses data from Cohort A of the Youth in Transition Survey (YITS-A) to provide some insight into this issue by examining the impact of school resources at the high school level on student academic performance at age 15 and on educational attainment at the post-secondary level. It thus supplements the contributions of Bedard (2003), Corak and Lauzon (2002, 2005) and Johnson (2005) to the somewhat limited empirical literature on the effect of school resources on educational outcomes in Canada. Seven academic performance measures – reported grades in math, science, and language, as well as the student's overall average grade and PISA scores in reading, math, and science, all at age 15 – and four alternative measures of educational attainment, two at age 19 and two at age 21, were examined. Ordered probit models of reported grades and educational attainment, as well as linear regression models of PISA scores, were examined. The results contain little evidence that such measures of school resources as student-teacher ratios, teacher quality, and the availability of computers and other physical and learning resources had much effect on reported grades, PISA scores, or educational attainment as of ages 19 and 21. However, school size and type of school – in particular, whether the school had a religious affiliation – had a statistically significant impact on PISA scores that was positive in the case of school size and negative in the case of religious affiliation. Attending a private school was also found to have a direct effect on some measures of educational attainment. These results suggest that more research is needed into the question of why private schools and schools with a religious affiliation performed differently from public schools with no religious affiliation.

Kathleen Day is a Professor at the University of Ottawa. (Kathleen.Day@uottawa.ca).

Table of Contents

Please cite as:

Day, Kathleen (2009). *The Effect of High School Resources on Investment in Post-Secondary Education in Canada. MESA Project Research Paper 2009-6*. Toronto, ON: Canadian Education Project. (www.mesa-project.org/research.php)

Acknowledgements

I would like to thank the Canada Millenium Scholarship Foundation and the Educational Policy Institute for research support through the MESA project, as well as Statistics Canada's COOL RDC for providing access to the data. I would also like to thank Theresa Qiu for assembling the data and doing much of the initial programming; without her assistance this paper would not have been completed. Finally, I thank the MESA Research Review Committee, Kim Foley, David Johnson, Ross Finnie, and other participants at the MESA sessions of the 2008 Canadian Economics Association Meetings for their helpful comments and suggestions.

Introduction

The promotion of investment in human capital has long been a preoccupation of Canadian governments. A well-educated and highly skilled work force is viewed as being essential to economic growth and prosperity. However, while few economists would dispute this view, there is less agreement on how best to improve the education and skill levels of Canadians. Should even more government funds be directed towards post-secondary education, or would directing more public resources towards the elementary and post-secondary levels be a better means of increasing the level of post-secondary educational attainment in Canada? Is public spending on education being used efficiently?

These are important policy questions that cannot be answered without hard evidence on the contribution of educational resources to educational outcomes. While a vast body of literature on the effect of school resources at the elementary and post-secondary levels on educational outcomes exists for the United States, there exist very few studies of this issue for Canada[1]. Bedard (2003) used Canadian Census data and quantile regression models to examine the effect of such measures of school resources as class size, school size, and teacher salaries on the earnings of Canadian men. Following Card and Kruger (1992), she examined the effect of provincial averages of these variables for earlier years on earnings in 1981, 1986, and 1991. Corak and Lauzon (2002, 2005) have also examined the role of class size and time-in-term, among other variables, using a semi-parametric approach applied to data from PISA 2000. However, these studies appear to be the only ones to date that have examined the direct effect of measures of school inputs on educational or earnings outcomes of Canadians[2].

Canadian research on the effects of school resources has undoubtedly been hampered by the lack of readily available data on both school characteristics and educational outcomes. However, in recent years, several new data sets amenable to this type of investigation have become available. Two of these data sets, both of which contain information on students and the characteristics of their schools, are the National Longitudinal Survey of Children and Youth and Cohort A of the Youth in Transition Survey (YITS-A). This paper uses data from the latter survey, which focuses on individuals aged 15 in 1999, to investigate the effect of various high school characteristics on several different outcome measures. These outcome measures include high school grades, standardized test scores from the OECD's Programme for International Student Assessment (PISA), and the level of educational attainment four to six years after the initial survey was administered. All the measures of school resources used pertain only to the school the respondent attended at age 15, since school characteristics were collected only in the initial year of the survey. Few measures of school resources consistently produced statistically significant effects

1 Hanushek (1996, 2006) has written several surveys of the literature on the effect of school resources and school quality on educational attainment and earnings. For other surveys of this literature, see Hedges, Laine, and Greenwald (1994), Hedges and Greenwald (1996), Betts (1996), Card and Krueger (1996), and Speakman and Welch (2006).

2 Studies other than the present one are known to be currently underway, but not yet published.

on the outcome measures examined. However, the type of school – private or public, with or without a religious affiliation – did appear to have an effect on the level of educational attainment in later years, as well as on PISA scores. Furthermore, some other measures of school resources appear to have been correlated with later educational attainment, despite their having little effect on academic performance at age 15.

The next section of this paper briefly reviews the previous literature dealing with the effect of school resources. The following section provides an overview of the data and econometric models used in this paper. The next two sections report on the effects of high school resources on student grades and test scores at age 15, as well as educational attainment at ages 19 and 21. The final section of the paper summarizes the results and their implications, discussing directions for future research.

Previous Literature

Because the relevant literature is extensive and there already exist numerous surveys of the literature on the effects of school resources, this brief review focuses primarily on the methods used to address these issues in empirical analysis, rather than trying to provide an exhaustive survey. The first two subsections examine the outcome measures and measures of school resources typically used in these studies. The third subsection reviews the econometric methods and models used, including a discussion of the econometric problems that arise in trying to measure the relationship between school resources and various outcome measures. Finally, the last subsection provides a brief discussion of the findings of the existing literature on this topic, with an emphasis on studies published since 2000.

Outcome measures

A variety of different outcome measures have been used as the dependent variables in studies of the effects of school resources, with standardized test scores being one popular choice. For example, Akerhielm (1995) used test scores in math, English, science, and history for a sample of American students who were in Grade 8 in 1988. The tests were administered as part of the U.S. Department of Education's National Education Longitudinal Study. Similarly, Marcotte (2007) used test score data from the Maryland School Performance Assessment Program, a standardized testing program for students in grades 3, 5, and 8. Other studies, such as Betts and Morell (1999), have used actual student grades; in particular, Betts and Morell used data on the Grade Point Averages of undergraduates attending the University of California at San Diego. Still other studies, such as Dearden, Ferri, and Meghir (2002) and Heinesen and Graversen (2005), have focused on measures of educational attainment. For example, Heinesen and Graversen used Danish data to examine the effect of school resources on the probability of achieving an upper secondary qualification, while Dearden, Ferri, and Meghir used data from the longitudinal National Child Development Survey of Great Britain and selected the level of education attained as of age 33 as their dependent variable.

All the outcome measures discussed so far have related to educational performance at various points in the life cycle. Yet another branch of the literature has attempted to measure the impact of school resources at the elementary and secondary levels on labour market outcomes – primarily wage rates. Like studies that have used educational attainment as their dependent variable, these studies have searched for longer term effects of school resources, and therefore have tended to focus on older individuals. Card and Krueger (1992) is one of the most frequently cited studies of this type. More recent studies that have used earnings or wages as a dependent variable include Heckman, Layne-Farrar, and Todd (1996) and Dearden, Ferri, and Meghir (2002)[3]. Eide and Showalter (2005) used unemployment status, not earnings, as their dependent variable, thus filling a gap in this branch of the literature.

Not only the basic outcome variable, but also the level of aggregation, differs across studies. Although many studies have used individual (student) data, some have used aggregate data at the regional or school level. For example, Marcotte (2007) used a panel data set in which the cross-sectional units were schools, not students, and the dependent variable was the percentage of students in school *s* at time *t* who achieved a satisfactory or excellent result on the standardized test. Similarly, Johnson (2005) used standardized test scores for Ontario schools in his analysis of the factors underlying school performance.

Measures of school resources

Just as there are many possible measures of student outcomes, there are also many possible measures of school quality or resources. Which measures are used by a particular study depends largely on the availability of data. When individual data sets are used in empirical analyses, data limitations also largely determine whether the school resource measures employed are specific to the school attended by the individual, or rather represent averages for broader regions such as school districts, states, or provinces.

One of the most popular measures of school resources is the student-teacher ratio or class size (e.g., Akerhielm 1995; Wilson 2001; Dearden, Ferri, and Meghir 2002). Other variables that have been used include measures of teacher qualifications, such as the proportion of teachers with a graduate degree (Eide and Showalter 2005; Rivkin, Hanushek and Kain 2005); teachers' test scores (Clotfelter, Ladd, and Vigdor 2007); and teachers' years of experience (Rivkin, Hanushek and Kain 2005; Clotfelter, Ladd, and Vigdor 2007). Other studies have focused instead on expenditure-related measures such as spending per pupil (Wilson 2001; Heinesen and Graversen 2005; Eide and Showalter 2005) and teacher wages (Bedard 2003). Both Heinesen and Graversen (2005) and Marcotte (2007) also included variables related to instruction time.

Locating data on these and other measures of school resources can be a formidable

3 As Hanushek (2000, 869) noted, the effect of school resources on labour market outcomes can be viewed as the result of a two-stage process: first, school resources influence the acquisition of cognitive skills, and then the individual's cognitive skills determine their labour market outcome. Therefore, studies that examine the effect of cognitive skills, as measured by standardized tests, on labour market outcomes are also related to this literature. Examples of such studies include Murnane et al. (2001) for the United States and Finnie and Meng (2002) and Green and Riddell (2003) for Canada.

task. For some researchers, such as Dearden, Ferri, and Meghir (2002) and Eide and Showalter (2005), the task was simplified by the fact that the survey data sets that they used included school questionnaires that provided information on the respondents' schools. In other cases, researchers using microdata had to merge data from several sources (e.g., Wilson 2001) and/or rely on region-level measures of school resources (Card and Krueger 1992; Bedard 2003). Needless to say, studies that rely on region-level data on school resources suffer from the limitation that these measures may not be very accurate for all individuals in the sample. In their study of the effect of school resources on earnings, Card and Kruger (1992) related the earnings of American adults to average school resources in their state of birth during the time they would have been in school. The implicit assumption underlying this approach was that individuals had obtained all their schooling up to a certain age within their state of birth. This approach was necessitated by the fact that they used Census data sets that did not include any information on school characteristics. Bedard (2003) applied similar methods to Canadian Census data.

As Hanushek (1996, 53) highlighted, because of variations between teachers and courses, even two students in the same school may not benefit from the same level of instructional resources; in addition, the resources available to the student can change over time as the student changes schools or educational policy in the student's school district changes. While some studies using aggregate data may contain information on changes in school resources over a period of time, few studies using individual data have had access to information on the school resources available to the individual throughout his or her education.

Although most studies have focused on the school resource measures listed above, a few have included other variables related to school type. For example, Dearden, Ferri, and Meghir (2002) included dummy variables indicating whether or not the secondary school attended by the respondent was private and/or co-educational. Because they also included a wide range of family background variables in their econometric models, they argued that this method of controlling for sample selection issues related to school type was preferable to an alternate approach excluding private school students from their sample.

Econometric issues

Most studies that examine the effect of school resources on educational outcomes are based on the concept of the education production function. This function is assumed to produce "education" in the form of test scores or educational attainment levels, with the inputs consisting of student inputs, family inputs, peer/neighbourhood influences, and school resources. Econometrically, the form of this function depends on the nature of the dependent variable used. If the dependent variable is a test score or numerical grade, linear regression models are estimated, as in Akerhielm (1995) and Betts and Morell (1999). If the dependent variable is related to the level of educational attainment, discrete choice models are typically estimated, as in Wilson (2001); Dearden, Ferri, and Meghir (2002); and Heinesen and Graversen (2005). Eide and Showalter (2005) also used discrete choice models to investigate the effect of school resources on the probability of being unemployed. Finally, studies of the impact of school resources on earnings typically estimate linear regression models that resemble human capital earnings equations augmented by the addition of school resource variables.

Regardless of the specifications of the dependent variables, and hence the econometric models to be estimated, there are certain problems faced by all researchers in this area. One of the most serious of these problems is the possibility that school resources may be endogenous. In other words, parents may have taken both school characteristics and their children's characteristics into account before choosing a residential location or a school, in which case the school resource variables may be correlated with the error term in the empirical education production function. Similarly, in her study of the effect of class size on elementary student test scores, Akerhielm (1995) argued that schools may take student ability into account when allocating students to classes of different sizes. In the absence of data on student abilities, such rules for allocating students to classes may result in a correlation between class size and the error term in the estimating equation.

Yet another source of endogeneity may arise from the manner in which school boards allocate resources to schools. For example, public schools located in what are perceived to be "high needs" neighbourhoods may be allocated higher levels of resources. At the same time, such schools may have more difficulty attracting good teachers. Such decision-making policies may thus strengthen the correlation between the school's location and school resources.

Endogeneity problems in models of educational outcomes may also be associated with other explanatory variables besides school resources. As Nechyba (2006) highlighted, researchers trying to measure the effect of peer influences on a student's performance also face serious endogeneity problems. The very notion that a student is influenced by his or her peer group implies a bi-directional relationship between the student's attitudes and those of the peer group – if student A is influenced by his or her peer group, then one must also allow for the possibility that student A influences the peer group. This problem is referred to in the literature as the "reflection problem."

These endogeneity problems have been dealt with in different ways by different researchers. For example, Akerhielm (1995) used two-stage least squares estimation, using school-level average class size and enrolment in Grade 8 as instruments for the student's actual class size. However, Dearden, Ferri, and Meghir (2002) eschewed instrumental variables methods in favour of simply including as many variables as possible to account for individual characteristics, family background, and local neighbourhood characteristics, since, as they pointed out, it is extremely difficult to find variables that explain the allocation of students to schools but do not also have a direct impact on student performance.

In studies of the effect of school resources on labour market outcomes later in life, researchers must contend with an additional problem: there is more than one channel through which school resources may affect outcomes. Heckman, Layne-Farrar, and Todd (1996) suggested that there are three such channels: school resources may increase the level of wages at all levels of education; they may affect the rate of return to education; and they may increase educational attainment (i.e., encourage individuals to remain longer in school). The latter two channels imply nonlinearities in the relationship between the log of earnings (the usual dependent variable in such studies) and measures of school resources. Heckman, Layne-Farrar, and Todd (1996) experimented with different methods of introducing nonlinearities into their model, which was based on that of Card and Krueger (1992).

Another type of nonlinearity that may exist is that the effects of school resources may be different for different sub-groups of students. For example, low-ability students may gain more from small class sizes than high-ability students (or the reverse). To test for such differences, one might estimate separate models for each group of students. Alternatively, some authors have tried instead to examine the entire distribution of outcomes, rather than focusing on the mean outcome as most studies do. Bedard (2003) and Rangvid (2007) did so using quantile regression methods, while Corak and Lauzon (2002, 2005) used semi-parametric methods instead.

Findings

A number of extensive surveys of the literature on the effects of school resources exist, with some of the most recent being those by Hanushek (2006), Hanushek and Rivkin (2006), and Speakman and Welch (2006). As in his earlier 1996 review, Hanushek (2006, 892) computed the proportion of studies that obtained statistically significant results with respect to the effects of two widely-used measures of school resources: the teacher-pupil ratio and expenditure per pupil. He found that 72 percent of 276 studies obtained an insignificant coefficient for the teacher-pupil ratio, while 66 percent of 163 studies that examined the effects of expenditure per pupil found that it too had an insignificant coefficient. Furthermore, in the case of the teacher-pupil ratio, only half of the studies that obtained a significant coefficient for this variable found that a higher teacher-pupil ratio led to improved student performance. In other words, in Hanushek's view, the empirical evidence on the effect of school resources was mixed at best.

One could argue, as did Hedges, Laine, and Greenwald (1994) and Hedges and Greenwald (1996), that simply counting the number of statistically significant results, without taking into account the type of data and methods used, does not constitute the best way to evaluate the literature on this topic. That said, an attempt to re-evaluate the vast literature on this topic would be beyond the scope of this study. Instead, it could be instructive to look at the results of a few recent studies not reviewed by Hanushek that have influenced the approach taken in this paper, including the few that have used Canadian data.

First, both Dearden, Ferri, and Meghir (2002) and Heinesen and Graversen (2005) examined educational attainment, with the former using British microdata and the latter using Danish microdata. Dearden, Ferri and Meghir (2002) estimated ordered probit models of educational attainment for both men and women at age 33, where the dependent variable consisted of seven categories ranging from “no qualification” to “university degree.” The pupil-teacher ratio was the only measure of school resources included in their model, although they did also include several dummy variables related to the type of school attended. They found that, although the pupil-teacher ratio had a large, statistically significant effect on male and female educational attainment in their simplest specification, this effect disappeared when school type and standardised test scores at ages 7 and 11 were added into the model. On the other hand, attending either a grammar school or a private school had a positive and statistically significant effect on the probability of achieving a high level of educational attainment[4].

In contrast, Heinesen and Graversen focused on only one level of educational attainment – the completion of upper secondary or vocational education – and used municipal-level school resource variables. In total, they were able to include four different measures of school resources, including expenditure per pupil and teacher wage hours per pupil per week. They also estimated four different types of models: logit models (with and without random effects) and linear probability models (with and without random effects). Like Dearden, Ferri, and Meghir, they found that their results with respect to school resources were sensitive to the number of control variables that were included. In their case, the effect of their principal school resource variable, expenditure per pupil, was statistically significant only when a full range of control variables (for personal characteristics, family background, and municipality socio-economic characteristics) were included in the model. In addition, the predicted effect of a 10 percent increase in expenditure per pupil was only about 1 percentage point, which represents quite a small effect.

Second, Rangvid (2007) studied the determinants of PISA scores in reading, math, and science for Danish students who participated in PISA 2000. Although her main focus was on the effect of the socioeconomic composition of the student body, she also included a number of school characteristics in her model, including the student-teacher ratio, school size, school autonomy, teacher morale, and several other variables. In both her OLS and quantile regression results, the school variables that appeared to have had the most important impacts on PISA scores were school size, school size squared, and an index of student-teacher relations. In general, all three had positive effects on student performance, although in the quantile regression models this effect was not statistically significant at all percentiles of the distribution of scores.

Turning now to the three Canadian studies that used Canadian microdata, as noted above, Bedard (2003) used Census data and hence had to use provincial average measures of school characteristics. Consequently, her school resources measures did not perfectly match the individuals in her data set. However, her school data did cover a relatively long period (1932-1952), since she pooled Census data for men from the 1981, 1986, and 1991 Censuses. Like Rangvid (2007), she too used quantile regressions, but her dependent variable was weekly earnings. The school resource measures included were class size, school size, and relative teacher wages. She found that, although class size did not have a statistically significant impact on earnings at the mean (which was the effect that would have been estimated in a standard linear model), it did have an impact at the extremes of the earnings distribution. Higher class sizes appeared to reduce earnings at the low end of the earnings distribution and raise them at the high end. The effects of school size and teacher salaries also varied across the income distribution.

The two studies by Corak and Lauzon (2002, 2005) are in many ways the most closely related to the present study, as they too made use of data from the YITS-A. However, the approach they used is quite different from that of the present study, in that they applied the semi-parametric approach of DiNardo, Fortin, and Lemieux (1996) to disentangle the effects of school factors and student backgrounds on the entire distribution of PISA scores. In addition, they included fewer control variables and fewer school characteristics than the present study. In their 2002 paper, they focused on the overall effects of school factors, rather than singling out individual measures, and, like Rangvid (2007), found that they differed at different points of the distribution of test scores. In their 2005 paper, they showed that if provincial differences in class size and time-in-term (i.e., the amount of instructional time) were eliminated by bringing all provinces to the Alberta levels, the distribution of student achievement in the Atlantic provinces would be significantly altered. However, once again, they found that the effect differed at different points in the distribution of test scores; for example, in the case of reading scores in New Brunswick, they concluded that the change would likely improve the average test score, but the proportion of students at the lowest level of reading proficiency would increase.

Overall, this very brief review of some relatively recent studies of the effects of school resources highlights the sensitivity of the results obtained to the specification of the model estimated. It also confirms that school resources do not affect all students equally, which suggests that a failure to observe a statistically significant effect of school resources in a model of mean achievement or educational attainment would not necessarily imply that school resources had no effect at all. While once again focusing on average outcomes rather than trying to look at the distribution of achievement, the present paper extends the limited Canadian literature on the topic by looking at more measures of educational achievement than Bedard (2003) and Corak and Lauzon (2002, 2005) – eleven in total – and by broadening the set of control variables to include more individual and family background variables, as well as some measures of neighbourhood and peer influences. Finally, more measures of school resources were included. The next section of the paper discusses the specification of each of the models estimated herein.

Overview of Data and Models

In this section, both the data used and the econometric models estimated in the present study are discussed. In the first subsection, a brief overview of the YITS-A data set is provided. The next subsection discusses the measures of school resources selected for inclusion in the empirical models. The third subsection describes the econometric models and the dependent variables used; in each case, the nature of the dependent variable determined the type of model that was estimated. The fourth subsection discusses the variables that were included to control for individual characteristics, family background, and neighbourhood effects. Finally, the last subsection provides a brief overview of the population examined.

The YITS-A data set

Cohort A of the YITS consisted of 29,687 Canadian 15 year-olds who were first interviewed in the spring of 2000. The respondents were re-interviewed every two years thereafter; to date, four cycles of data on these individuals have been released. In the initial year of the survey, not only students but also their parents and school principals were interviewed. Both public and private schools were included in the survey. In addition, all students participating in the survey were administered the OECD's PISA test of reading skills in the initial year. Roughly half the students also wrote the PISA math test, while the other half wrote the PISA science test. The combination of all these elements made the YITS-A data an extremely rich source that could be used to analyze a wide variety of education-related issues.

For the purposes of this study, it was decided to include in the sample for each model all respondents for whom all the necessary data were available. Because the different outcome variables selected for analysis were not all available for all the students, the sample sizes therefore varied for the different models presented, ranging from 6,981 to 12,162. While this approach somewhat reduced the comparability of the results across models, it did ensure that the sample sizes were fairly large.

Measures of high school resources

The PISA/YITS School Administrator File contains a great deal of information about the high school attended by each respondent at age 15, ranging from pupil-teacher ratios to an index of teacher morale. While all of these variables may be somewhat subjective in that they reflect the school principal's impressions, some could be more readily linked to expenditure or regulatory policies related to education. These less subjective variables were the measures considered to be of greatest interest to this study.

First of all, following the example of Dearden, Ferri, and Meghir (2002), several dummy variables relating to the type of school were included: a dummy variable equal to one if the school was co-educational and zero otherwise[5], a dummy variable equal to one if the school was private and zero otherwise, and a dummy variable equal to one if the school had a religious affiliation and zero otherwise. The inclusion of these dummy variables tested whether there existed any

5 Note that in most Canadian provinces there exist two parallel publicly-funded educational systems, one secular and one associated with the Catholic Church. There also exist some private schools associated with other religious denominations.

differences in educational performance associated with these particular types of schools.

Second, five variables related to the educational resources of the school were included: the total number of students in the school; the student-teacher ratio; the ratio of computers to students in the school as a whole; an index of the quality of the school's physical resources (e.g., space, lighting, condition of building); and an index of the quality of the educational resources (e.g., library materials, computers, lab equipment) available to students of the school[6]. The first two of these variables allowed for the possibility that the absolute size of the school or the average class size may have affected the student's performance; the third variable was included to test whether there would be any evidence that access to computers at school was an important factor in academic performance. The two index variables, although somewhat subjective because they were based on the opinions of the school principal, were included because both physical and educational resources are variables related to the expenditures required to operate the school. These could also be thought of as measures of two types of capital inputs into the production of education.

Next, at least one measure of the quality of the teaching staff was included in each model. In total, four such measures were available: the proportion of the teaching staff with a bachelor's degree in education; the proportion of math teachers with a bachelor's degree in math; the proportion of language and literature teachers with a bachelor's degree in that subject; and the proportion of science teachers with a bachelor's degree in science. Many high school teachers found themselves teaching in more than one field; furthermore, school principals may not always have been able to guarantee that all teachers assigned to teach a particular subject had a teaching qualification in that subject. The inclusion of these variables tested whether the educational qualifications of teachers had a measurable impact on student performance. In models in which the dependent variable related to a particular subject (for example, math), only the teacher quality variable that related to that subject was included.

It should be noted that the PISA School Administrator questionnaire also included a number of questions about the availability of special programs or courses. One variable related to these types of programs was included: a dummy variable equal to one if the school offered special courses in study skills and zero otherwise. The inclusion of this variable tested whether the availability of such courses had an impact on student performance. However, no variable was included to indicate whether the respondent actually took such a course.

In addition to all these variables, three variables related to the teaching climate were included: an index of teacher morale; an index of school autonomy; and an index of teacher participation. All three indexes were constructed by the OECD (for details regard-

6 The index of the quality of physical resources and the index of the quality of educational resources were constructed by the OECD from responses to questions about the adequacy of the physical and educational resources of the school. See Chapter 17 of OECD (2003) for further details on the construction of these indices.

ing their construction, see OECD 2003). Although these variables were not necessarily directly related to the school's financial resources and were also somewhat more subjective than the other variables discussed so far, these final variables were added to the model after the initial estimation results suggested that attending a private school was strongly positively related to academic performance, because it was thought that they might help to capture some of the differences between private and public schools.

The set of variables relating to school characteristics and resources included in the present analysis is quite large in comparison to those included in other studies. Although small in relation to the number of variables available in the School Administrator File, it constituted a compromise between the desire to include as many characteristics as possible and the need to keep the specification parsimonious. Furthermore, most of the chosen variables could be directly related to educational policy. For example, a finding that teacher qualifications mattered would suggest that governments place a higher priority on ensuring that all high school teachers be properly qualified.

Although this list of school characteristics was believed to be appropriate, the possibility of multicollinearity between the various measures suggests that it might have been worth trying to combine the measures into one or more indices. However, the difficulty with such an approach is that it was not obvious what weights should be applied to the different measures. As an alternative to arbitrarily assigning weights to the various measures, principal components analysis was used in an attempt to reduce the set of school characteristics to a considerably smaller number of explanatory variables. However, because the measures were not in fact as highly correlated as one might have expected, this approach did not prove to be very successful.

In total, nineteen variables (including school climate variables related to the impact of bullying, drug use, and poor home environments, all discussed later) were included in the principal components analysis. However, at least fifteen of the nineteen principal components were required to account for 90 percent of the total variation in the set of nineteen school characteristics. Given the general difficulty of interpreting principal components, the gain in parsimony did not seem large enough to warrant replacing the nineteen original variables with the fifteen most important principal components. Therefore, this attempt to produce a small number of indices of school characteristics was deemed a failure[7].

Finally, in the educational attainment models, one additional variable related to school resources was included: the PISA reading score at age 15. This variable was used as a general measure of cognitive skills, which were presumed to be one input into the education production function later in life. If high school characteristics had affected PISA scores, and PISA scores influ-

7 Principal components analysis was also applied to several subsets of school characteristics, such as the measures of teacher quality. However, once again it did not seem to be possible to construct a single index that captured a large proportion of the total variation in any of the subsets of characteristics examined.

enced educational attainment, then school characteristics might have had an indirect effect on educational attainment through the PISA score. Although high school grades were also available, they were not included in the models of educational attainment for reasons that are discussed in the next subsection.

Outcome measures and econometric models

Although the YITS-A data set contains a large number of variables related to school characteristics, all of these variables pertain only to the school attended at age 15. No new information about school characteristics was collected after cycle one. While this lack of historical information on school characteristics was not unique to this study, it did pose a problem for the estimation of the effects of school characteristics on educational attainment in later years, since the school characteristics may have changed over time either because the student changed schools or because of changes in school funding or administrative policies. To deal with this problem, the empirical analysis was divided into two parts.

First, the relationship between school characteristics and academic performance at age 15 was examined. Then, educational attainment as of ages 19 and 21 was examined. In the first stage of the analysis, both high school grades and PISA scores were used as dependent variables, while in the second stage the PISA reading score was included as an explanatory variable in models of educational attainment.

The YITS-A data set contains several measures of academic performance as of age 15. Several grades reported by the student at age 15 were available: the overall average in school in 1999, the mark in the most recent math course, the grade in the most recent science course, and the mark in the most recent "language" course[8]. All four were used as dependent variables in the first part of the econometric analysis. Since these grades were reported as categorical variables, ordered probit models were used[9].

The categorical variables provided in the original data set were transformed so as to increase in value as the grade increased as follows: 1 was associated with a grade of less than 50 percent; 2 was associated with a grade from 50 percent to 59 percent; 3 was associated with a grade from 60 percent to 69 percent; 4 was associated with a grade from 70 percent to 79 percent; 5 was associated with a grade from 80 percent to 89 percent; and 6 was associated with a grade from 90 percent to 100 percent. In addition, the student's PISA scores for reading ability, math and science were available and were also used as outcome measures[10]. Because they were derived from standardized tests, the PISA scores may have provided more accurate measurements of student ability than

8 A "language" course here does not mean a course in a second language, but rather a literature course. The term "language course" was used in the YITS-A documentation because the language of instruction may have been either English or French, depending on the school system in which the student was registered.

9 See Cameron and Trivedi (2005, 519-520) for a brief overview of the ordered probit and logit models.

10 To be more specific, the weighted likelihood estimates of each student's score were used. According to OECD (2002), these are the best measures to use in analyses of individual student performance. The weighted likelihood estimates are estimates of student proficiency derived from the student's responses to the PISA tests. The PISA assessment file also contains alternative measures for each student, called *plausible values*. See OECD (2003) for more details.

recorded grades, which may have varied across schools for a variety of reasons. Since the PISA scores were not categorical variables, linear regression methods were used to analyze them.

In the second stage of the analysis, educational attainment as of cycle three (at age 19) and cycle four (at age 21) was examined. Like grades, educational attainment was reported as a categorical variable, rather than as total years of schooling. Furthermore, in each cycle, two educational attainment variables were available. The first consisted of four categories: has not yet graduated from high school; is a high school graduate, but has no post-secondary education (PSE); has some PSE, but has not yet graduated from a post-secondary program; and has graduated from a post-secondary program[11]. Because there was a clear hierarchy in the ordering of the educational categories here, ordered probit models were estimated for this variable.

While it would be of interest simply to know whether or not an individual student had attempted or completed PSE, it would be more ideal to have more information about the type of PSE undertaken. Fortunately the second educational attainment variable in the YITS-A data set provided more information about the type of PSE undertaken, dividing it into 14 categories[12]. However, because the respondents were only 19 and 21 at the time of data collection in cycle one and cycle two respectively, only a small number of respondents had enrolled in certain types of PSE (yet). For this reason, the 14 categories were aggregated into ten, leading to the creation of a new dependent variable with 11 categories.

The first category consisted of individuals who had not yet taken any PSE; this category included both high school graduates and those who had not completed high school. The second category included three levels of PSE below the college or CEGEP level. Each of the next seven categories corresponded to the next seven categories of the original variable, ending with the bachelor's degree. Category ten included the Master's and Ph.D levels, and the final category included the last two categories of the original variable.

This second measure of educational attainment thus provided much more detailed

11 These are the YITS-A variables HEDLD3 (Cycle 3) and HEDLD4 (Cycle 4).

12 These are the YITS-A variables HLPSD3 and HLPSD4, which were defined as the "highest level of post-secondary education taken across all programs and institutions as of December 2003 (or 2005)." Listed in the order given, the categories included Attestation of Vocational Specialisation (AVS or ASP); Private business school or training institute diploma or certificate; Registered Apprenticeship program; College or CEGEP program; University transfer program at a college or CEGEP (for credits, university transfer diploma or Associate's Degree); College post-diploma or graduate level program (college diploma or higher needed first; University diploma or certificate BELOW Bachelor's (undergraduate level); Bachelor's degree; First professional degree; Graduate-level diploma or certificate ABOVE Bachelor's, BELOW Master's; Master's degree; Ph.D. degree; Diploma, certificate or license from a professional association as in accounting, banking, or insurance; and Other level of post-secondary.

information about the type of PSE attempted. Because it too was an ordered variable, an ordered probit model was also used to analyze this measure of educational attainment[13].

Control variables

As Dearden, Ferri, and Meghir (2002) argued, in order to avoid bias in the estimation of the coefficients of school resource variables, it is important to include as many variables as possible to control for individual characteristics, family background, and neighbourhood or peer effects. In all of the models estimated in this paper, the same control variables were used. First, dummy variables indicating whether or not the individual was female, belonged to a visible minority group and/or was an immigrant were included. In addition, a variable measuring years since immigrating to Canada was included to allow for the possibility that immigrants become better integrated into the school system the more time they spend in Canada.

To allow for the possibility that disabilities or health problems might affect academic achievement, four dummy variables were included to control for limitations of this nature. These variables were constructed by combining information from a YITS-A question to parents about whether their child had a physical condition, mental condition, or health condition that interfered with their child's schooling, as well as a series of questions relating to the type of condition[14]. The four variables indicated whether the child had a physical disability, a physical health problem, a learning disability, or some other condition that interfered with his or her schooling.

Other student characteristics included indicators of the frequency of smoking and a series of dummy variables measuring the number of schools attended up to age 15. Smokers were divided into those who smoked at least once a day and those who smoked no more than once or twice a week. Smoking behaviour was included because it may be an indicator of the student's rate of time preference[15]. The number of schools attended was included because frequent changes in schools may be disruptive and may have a negative impact on the student's academic performance. In the models of educational attainment an additional variable

13 The ordering of the categories of this variable followed that used by Statistics Canada in constructing the original variables HLPSD3 and HLPSD4. Although one could question the validity of the ordering (does a university certificate below the bachelor's level really constitute a lesser educational qualification than a college graduate-level program?), it no doubt corresponds to a generally-agreed upon ranking of such qualifications. Furthermore, given the large number of categories, it would have been extremely difficult to estimate a less restrictive multinomial probit model of these dependent variables. Finally, although multinomial logit estimation would have been possible, it would have depended on the assumption of independence of irrelevant alternatives. When attempted, it also yielded somewhat odd results for some coefficients, not to mention a very large volume of output.

14 The YITS-A variables used were PD10B and PD1101-PD1111, all of which were included in the parent file. The four dummy variables were defined to be equal to one if the child had a condition that sometimes or frequently affected the activities they could do at school. Hearing problems, vision problems, speech problems, orthopaedic problems, and other physical disabilities were all included in the physical disabilities variable; physical health problems included allergies, asthma, and diabetes; and "other disabilities" included mental health conditions as well as other undefined conditions.

15 The rate of time preference is the rate at which the individual must be compensated for forgoing say \$100 of consumption today in return for consumption in the next period. If the individual's rate of time preference is ρ, they will require a future payment of at least \$100(1+ ρ) in order to induce them to give up \$100 today. Smokers are believed to have a higher rate of time preference than non-smokers, since appear to care more about their current enjoyment of tobacco products than about the possible future health consequences of smoking. I thank Kelly Foley for suggesting the inclusion of these variables and their interpretation.

measuring the number of high schools attended was also included.

Family background was represented by a total of twenty-five variables. The first four were dummy variables representing the number of siblings. Because birth order has been cited by some authors as a determinant of individual outcomes, four additional variables controlling for the number of older siblings were also included. The ninth variable measured the number of times the student had moved; because changes of household can occur without a change of schools, this variable was not perfectly correlated with the variables reflecting the number of schools attended. An index of family educational support, based on the student's responses to questions about how often other family members helped the student with school work, was included. It was assumed that families that were more willing to help with school work may also have placed a high value on education. Family income as of cycle one, as well as three other variables representing family structure as of cycle one, were also included.

In addition, four dummy variable indicators of parents' aspirations regarding the level of education they hoped their child would achieve were included, because differences in parents' aspirations for their children might translate into differing degrees of pressure on the student to perform well in school. The first of these variables was equal to one if the student's parents hoped that their child would obtain a trade certificate or vocational qualification; the second was equal to one if they hoped that their child would attend college; the third was equal to one if they hoped that their child would obtain a university degree; and the last was equal to one if they hoped that their child would obtain a post-secondary qualification, but didn't care what kind.

The final eight family background variables were indicators of the parents' level of education. The omitted category among this group of variables was having at least one parent who graduated from high school, but no parents with post-secondary education. Another dummy variable indicated that neither parent finished high school. The remaining six dummy variables reflected various levels of post-secondary education.

Although previous studies have found peer influences to be important, there has been some disagreement among researchers with regard to how to define the peer group. Some authors, such as Solon, Page, and Duncan (2000) and Ludwig, Ladd, and Duncan (2001), have defined the peer group as children living in the same neighbourhood. Others, such as Hanushek, Kain, Markman, and Rivkin (2003), have defined the peer group as students in the same grade at the same school, and have used the proportion that were black, the proportion that were Hispanic, and the proportion that were eligible for a free lunch as their peer variables. While these two definitions were likely to overlap in many cases, in some cases – particularly for those students who attended private schools – they may not have been perfectly correlated.

For the purposes of this study, it was deemed desirable to include both some neighbourhood characteristics and some characteristics of the student's friends and classmates. Unfortunately, the YITS-A data was somewhat limited with regard to neighbourhood characteristics; the only relevant information available pertained to the province in which the student's school was located and the population size of the community in which the school was located.

While linking the YITS-A data to Census data on neighbourhoods might have been possible, it would have required more detailed information from Statistics Canada regarding the Census district in which the student lived or where his or her school was located. However, because a school-based sampling scheme was used to construct the YITS-A, it was possible to construct school-population characteristics from the YITS-A data itself. More specifically, three variables reflecting the characteristics of schoolmates' families were constructed by averaging data from respondents from the same school on family income, parental employment status, and whether or not either parent had a university degree[16].

To ensure that these variables measuring average family income, the proportion of families with at least one parent employed, and the proportion of families in which at least one parent had a university degree were of reasonably good quality, respondents from schools with less than twenty students included in the YITS-A were eliminated. The other neighbourhood characteristics consisted of nine provincial dummy variables (Ontario was chosen as the reference province) and five community size dummy variables. This set of neighbourhood characteristic variables was much smaller than that employed by Dearden, Ferri, and Meghir (2002) and Heinesen and Graversen (2005), but it was the best that could be done with the available information.

Although it does not contain much information regarding neighbourhoods, the YITS-A data set does contain a fair amount of information about peer influences, with peers being broadly defined to include both siblings and friends. Among other things, respondents were asked about their friends' attitudes towards homework and post-secondary education.

Unfortunately, in a study of school resources, it is impossible to employ one of the methods recommended for avoiding the reflection problem discussed by Nechyba (2006), namely including school fixed effects, since doing so would eliminate all school-based variables from the analysis. In part for this reason, the number of peer-related variables included was limited to six. Three of these variables were derived from student responses to a question about what proportion of the student's friends were considered to be trouble-makers. The other three were derived from the school administrator questionnaire, which included questions relating to the extent to which the principal thought that learning at the school was inhibited by bullying, drug and alcohol use, and the poor home environments of students.

For each of these latter three factors, a dummy variable was defined to be equal to one if the principal's response was "to some

16 I am grateful to David Johnson for suggesting this approach.

extent" or "a lot." Although their accuracy would have been dependent on how well the principal knew his or her own school, these three variables did provide some information about the school climate and peer influences operating on the students attending that school. Moreover, because these variables were derived from the school administrator's questionnaire rather than from student responses, they may have been less likely to cause endogeneity problems than other peer-related variables.

In light of the large number of explanatory variables included in the model, a complete list of variable names and definitions is provided in Table 1. The variable names included in the table are identical to those used in the subsequent tables.

An overview of the sampled population

Because different samples were used in the estimation of each model, descriptive statistics were provided for only one of the samples used as the project progressed. All the descriptive statistics presented in Tables 2 through 7 comprise population values constructed using sampling weights. They thus provide an overview of the characteristics of the population underlying the research.

Table 2 contains a breakdown of the sampled population by school type, as of cycle one. This table shows that approximately 92 percent of the population of 15-year-olds in 1999 attended public schools. About 31 percent of these public schools had some religious affiliation. In contrast, the majority of private school students – about 83 percent – attended schools with a religious affiliation.

Of particular interest to this study is the relationship between school characteristics and academic performance. Table 3 shows a frequency distribution of students by school type and the overall average grade reported by the student at age 15. Overall, approximately 9 percent of students performed poorly, in that they received overall average grades of 59 percent or less. This proportion was considerably lower among students in private schools; less than 3 percent of private school students received grades below 59 percent. At the other end of the distribution, private school students were also more likely to have received a grade above 90 percent – 9 percent of private school students, as compared to 7.4 percent of public school students, fell into this category.

Because PISA scores were the outcome of standardized testing, it was interesting to compare the distribution of PISA reading scores in Table 4 to the distribution of overall average grades in Table 3. For this purpose, the PISA reading levels specified by the OECD were used. For both public and private school students, the proportion who achieved no better than Level 1 was identical to the proportion who received overall average grades of 59 percent or less. However, the proportion of students at Level 5 – 24 percent for private school students, 15 percent for public school students – was considerably higher than the proportion receiving overall average grades of 90 percent or better. This observation suggests that, although literacy may have been essential for good academic performance, exceptionally strong reading skills did not guarantee that students would receive exceptionally good marks. In contrast, good math and science skills seem to have been more important for obtaining exceptionally high average grades.

Although the comparisons made in Table 3 and Table 4 provided some support for the popular notion that private schools may provide a better quality education than public schools, they did not reveal anything about the relationship between school resources and academic performance. Table 5 and Table 6 contain calculations made in an attempt to fill this gap by showing the means of eight school resource variables that were closely related to educational funding for each level of academic performance. The first feature of these tables to note is that there was surprisingly little variation across levels of performance in many of the measures.

As shown in Table 5, standard measures of school resources, such as the student-teacher ratio, the proportion of qualified teachers, and the ratio of computers to students, were almost identical across grade levels; with the exception of the lowest two levels of PISA scores, the same pattern can be seen in Table 6. It is interesting to note that, while Table 6 suggests a positive correlation between PISA reading levels and school size (as measured by the number of students), no corresponding relationship with average grades is evident in Table 5.

Of the eight measures of school resources included in Table 5 and Table 6, the two that seemed to exhibit the greatest variation across levels of academic performance were the indexes of physical and educational resources available to students (*scmatbui* and *scmatedu*). However, the relationship between these variables and academic performance was difficult to discern from the tables. While, as seen in Table 5, the index of educational resources seemed to fall and then rise as grades increased, Table 6 shows that there seems to have been to be a fairly steady downward trend in this variable as PISA scores increased.

Before considering what the estimation results revealed about the factors influencing academic performance, it is helpful to take a brief look at some of the individual and family characteristics of the sample. Table 7 shows the means of many of the explanatory variables included in the model. For example, approximately 50 percent of the sampled population were female; about 7 percent were immigrants to Canada; and 11 percent belong to a visible minority group. Only 7 percent of the sampled population were only children, while more than 50 percent had at least one older sibling. The average family income of the sampled population was approximately $69,000 in 1999. Most of the respondents lived with their two biological parents. Only 11 percent lived in rural areas.

The Effect of High School Resources on Student Performance at Age 15

Turning now to the first stage of the analysis, Table 8 contains the coefficient estimates of ordered probit models of high school grades reported at age 15. Table 9 contains the coefficient estimates for the student's scores on the standardized PISA tests for reading, math, and science. The latter equations were estimated using Ordinary Least Squares (OLS). Both the Table 8 and Table 9 estimates were weighted using sampling weights. Robust standard errors, with a correction for clustering at the school level, were computed in both cases. In each case, the largest possible sample – i.e., the largest sample for which all the required variables were available – was used to estimate the model. Since individuals who reported certain grades did not always have PISA scores recorded, or vice versa, and the different measures of teacher qualifications were not always available for all students, the sample size was different for each model estimated. The sample sizes varied from 6,981 observations (for the PISA science score equation shown in Table 9) to 12,162 (for the PISA reading score equations, also shown in Table 9). The sample sizes were considerably smaller for the PISA math and science scores because these tests were not administered to all students, as was the PISA reading test.

In interpreting the estimation results, the identity of the reference student should be kept in mind. The reference student was male, Canadian-born, and not a member of a visible minority group. He had attended three different schools up to age 15, a number not atypical for the Canadian school system. He had no disabilities or health problems, and was a non-smoker. He had one sibling, but no older siblings. In terms of their educational aspirations for him, his parents did not hope for him to achieve any more education than a high school diploma. He lived with both biological parents. At least one of his parents had graduated from high school, but neither had any post-secondary education. None of his friends were considered trouble-makers. Finally, he lived in Ontario in an urban area that qualified as a city, but not a large city, and attended a school that was public, had no religious affiliation, was not co-educational, and did not offer any special courses in study skills.

If high school grading systems were uniform across schools, it might have been expected that the results of the two sets of estimates be quite similar, at least in terms of their qualitative implications. However, the correlation between the two measures of student proficiency was not as high as might have been expected. To facilitate comparisons between these two measures, the PISA reading scores were converted to a categorical variable with six possible values corresponding to the six levels of reading proficiency defined by the PISA project (OECD 2003, 207). Reading proficiency increased as the value of the categorical variable increased from 1 to 6.

To measure the correlation between PISA scores and reported grades, Spearman coefficients of the correlation between the categorical PISA reading level variable and reported grades were computed for a sample of 14,997 students. In each case, observations for which the relevant high school grade was unavailable were deleted from the calculation. Interestingly, the correlation between the PISA reading level variable and the last language grade was only 0.43 (based on

13,857 observations). The overall average grade showed the strongest correlation with the PISA reading level at 0.49 (13,814 observations). The correlations between the reported grades for different subjects were generally considerably higher, ranging from 0.48 (between the math grade and the language grade) to 0.76 (between the science grade and the overall average)[17].

These relatively low correlations could simply reflect the fact that the six reading levels defined by PISA did not correspond well to the grading levels used in Canadian high schools. The criteria used may simply have been very different in the two cases. Inaccurate or untruthful reporting by the students surveyed may also have been part of the problem. Nonetheless, the low correlations between the PISA scores and the reported grades did raise questions about grade inflation at the high school level, and whether or not high school grades might provide an accurate measure of student abilities.

Since high school grades and PISA scores were not as highly correlated as one might have expected, it would not have been surprising to find that there were some important differences between the estimation results for these two types of measures. Such differences did indeed exist. Nonetheless, there were quite a few important similarities between the two sets of results.

In comparing Table 8 and Table 9, it was apparent that many of the individual and family background variables that had a statistically significant effect on one set of estimates also had a statistically significant effect on the other. For example, sex (*fem*) and immigrant status (*imm*) had statistically significant coefficients in most of the seven models estimated. However, their effects were not always the same. While being female significantly increased both types of reading and science performance measures, the calculations in Table 8 showed that being female also raised math grades and overall grades, while the Table 9 estimates implied that girls were *less* proficient in math than boys, holding all else constant. According to Table 8 estimates, belonging to a visible minority group tended to be positively correlated with one's reported grades; along the same lines, according to Table 9 estimates, belonging to a visible minority group was consistently negatively correlated with lower levels of proficiency in all three domains tested.

The results with respect to immigrant status were also contradictory, with the PISA results suggesting that immigrants tended to perform worse than Canadian-born students in science and reading proficiency (although this negative effect tended to decrease as the number of years since migration to Canada increased), while the grade results suggested that immigrants did better in math and worse in reading. The finding that immigrants

17 The magnitude of the correlations between the different subject grades are not too surprising. There is certainly a perception on the part of many people that students who are strong in math are not necessarily strong in the arts, and the relatively low correlation between the reported math and language grades supports this view. The high correlation between the science grade and the overall average may reflect the fact that it is often easier to obtain high grades in math and science courses, where tests tend to be more objective, than in arts and humanities courses. The correlation between the overall average grade and the math grade was also relatively high at 0.70.

tended to perform worse than Canadian-born students in language courses or reading proficiency was not surprising, since a large proportion of recent immigrants to Canada have arrived from countries where neither English nor French is the primary language.

Not surprisingly, there was a strong and statistically significant negative relationship between learning disabilities and academic performance. Holding all else constant, on average, students with learning disabilities received PISA scores that were between 43 and 63 points lower than those of students without learning disabilities. The probability of receiving an overall average grade of at least 70 percent was also lowered by 0.26 for these students as compared to the base case probability of 0.58. In other words, the probability of obtaining an overall average grade of at least 70 percent dropped by almost 50 percent if a student had a learning disability. Other types of disabilities or health problems did not appear to have a significant effect on academic performance, however.

Other findings that were consistent across the two tables were the negative effects of smoking and having attended only one school. Smoking at least once a week had a strong negative impact on both PISA scores and high school grades in all subject areas. To the extent that smoking behaviour reflected a high rate of time preference, the negative effect observed may have been associated with an unwillingness to spend time studying. Having attended only one school was significantly negatively correlated with student performance, in comparison with students who had attended at least three schools by age 15. It was not clear why this should have been the case. Surprisingly, there was little evidence that attending a large number of schools was negatively correlated with student performance. However, changing households frequently did seem to have a negative and significant correlation with academic performance in all the models estimated.

In five of the seven models presented in these two tables, there was a significant negative correlation between having no siblings and academic performance. Having two siblings also had a significant negative coefficient in both tables, but only in the equations for science scores. However, it appeared to be the number of older siblings, rather than the total number of siblings, that had the most important effects on performance in both tables. All four older sibling variables had statistically significant coefficients in almost all of the seven models; the only exception was the coefficient of *oldsib1* in the math grade equation in Table 8. Moreover, as the number of older siblings increased, the magnitude of the negative coefficient tended to increase in most of the equations.

In other words, birth order had a fairly important effect on academic performance at age 15, with the eldest child in a family tending to perform best, holding all else constant. For example, in terms of PISA reading scores, having just one older sibling reduced the score by 8.75 points, while having four or more older siblings reduced it by 19.5 points, more than twice as much. Summing marginal effects across grade levels, having one older sibling raised the probability of an overall average grade below 70 percent by

0.02 (relative to a base probability for the reference student of 0.42), and lowered the probability of a grade higher than 69 percent by a corresponding amount. Having at least four older siblings raised the probability of having an overall average grade below 70 percent by 0.08.

Although it did seem to be important, the number of older siblings was not the only family background variable that had a statistically significant correlation with student performance. A somewhat surprising result was that family income (*faminc*) did not have as big an impact on student performance as might have been expected; it did not have a statistically significant coefficient in any of the seven models estimated and, furthermore, its estimated coefficient was small in magnitude. This finding could be viewed as being consistent with that of Frenette (2007), who found that only 12 percent of the gap in university attendance rates across the Canadian income distribution could be explained by financial constraints.

However, other family background variables had more important relationships with student performance. The higher the parents' aspirations with respect to the level of education their child would ultimately achieve, the higher the student's PISA scores and high school grades. Having parents who hoped for their child to get a university degree (*euni* = 1) raised PISA scores by 52 to 66 points at the 1 percent level of significance, depending on the domain, and also had a positive impact on all grades. The probability of receiving an overall grade above 69 percent also went up by approximately 0.14, relative to a base probability of 0.58.

Living in a two-parent household in which at most one parent was a biological parent also had a negative effect at the 1 percent level of significance in all seven models, a finding which may reflect the disruptive effects of changes in family structure on children. Interestingly, living with a single parent appeared to have less of a negative impact on academic performance; although it had a statistically significant coefficient in three of the four models for reported grades, its coefficient was smaller than that of the two-parent family dummy. Furthermore, the single-parent dummy variable did not have a significant coefficient in the PISA score models.

One family background variable for which the results seemed particularly surprising was the family educational support variable (*famedsup*). This variable, constructed by the OECD, increased with the amount of help the student received from parents and siblings. The negative and significant coefficient of this variable in six of the seven models, and particularly in all three PISA score models, suggests that such help was actually hindering the student. One can think of several possible reasons why this might be the case: for example, perhaps the variable was actually an indicator of lower ability on the part of the student, because students with high levels of ability would not need as much help, particularly by the time they had reached age 15. Alternatively, perhaps family members were helping too much, to the extent of sometimes doing the student's work for him or her. If this was the case, then students may not have been learning as much as they should have been, a factor that might have resulted in the observed lower PISA scores.

Far more important than family educational support, however, were the parental education variables. Having parents who did not graduate from high school (*nohs* = 1) was significantly negatively correlated with all three PISA scores, as well as with overall average grades and language grades. By way of contrast, having at least one parent with a college diploma (*diploma* = 1) or a certificate or degree from a university (*ucert, bach, prof,* or *grad* = 1) tended to be strongly positively correlated with academic performance, with the magnitude of this effect rising with the level of parental education. For example, students who had a parent with a college diploma had a PISA reading score that was about 13 points higher than that of otherwise identical students; if at least one parent had a graduate degree from a university, the child's PISA score was about 51 points higher on average, holding all else constant. In the case of parents who had a graduate degree, the child's probability of receiving an overall average grade above 69 percent also increased by 0.21, or about 36 percent of the base probability of 0.58.

Turning now to the "neighbourhood" effects, the size of the community in which the student's school was located did not have too many significant relationships with student performance. However, there were a number of significant provincial effects[18]. Notably, the pattern of the provincial effects was not always the same for the reported grades and the PISA scores. For example, in all four grade equations the coefficients of the provincial dummies were strongly positive for Prince Edward Island (PEI) and New Brunswick; however, they were generally strongly negative for these provinces in the PISA score equations (the one exception being the PISA math score equation, where the coefficient of the PEI dummy was not statistically significant). The reverse was true for Alberta: in the grade equations, the provincial dummies had significant negative coefficients, while in the PISA score equations they had significant positive coefficients.

These conflicting results with respect to the provincial dummies suggest differences in grading standards across provinces. If the PISA scores, which were derived from standardized tests, were taken to be the most accurate measures of student proficiency, then these results might suggest that the Alberta school system was doing a better job of training students than was the Ontario system. At the same time, it appears that Ontario grades were perhaps inflated in comparison to grades in Alberta. The grade inflation problem appears to have been even more serious in Prince Edward Island and New Brunswick, since their reported grades tended to be higher than those of Ontario, yet their students tended to receive lower PISA scores than did Ontario students. Table 9 shows that students in Quebec and Manitoba also tended to perform better on the PISA tests than did Ontario students after controlling for individual and family characteristics. The results for the coefficient of the Saskatchewan dummy suggest that students from that province performed slightly better than did Ontario students on the PISA math

18 It is possible that the significance of some of the provincial dummy coefficients might have changed if clustering at the provincial level had also been taken into account in computing the standard errors. Since education is provincially regulated in Canada, there might well be some correlations between schools within each province.

and reading tests, and received considerably higher grades than Ontario students.

The performance of the remaining "neighbourhood" variables, namely the average family income of students attending the same school (*avefaminc*) and the proportion of parents who had a university degree (*puni*) or who were employed (*ppemp*) was somewhat surprising. There appeared to be no relationship at all between the proportion of parents employed and academic performance, while average family income was significantly related only to the PISA math score. The coefficient of *puni* was statistically significant in four of the seven models, however, suggesting that there was some correlation between the education level of the parents of other students at one's school and one's academic performance. But although this correlation generally seemed to be positive, it was negative in the case of language grades in high school. It is not clear why this should have been the case.

Of the six variables included to reflect peer influences, those that seemed the most strongly correlated with academic performance were those indicating the proportion of the student's friends that were considered to be trouble-makers. If all or most of the student's friends were considered to be trouble-makers then the student's academic performance tended to be significantly worse. High school grades were even significantly lower for students who counted only a few trouble-makers among their friends; however, the reading score was the only PISA score that was significantly affected in this case, and the effect was not large. If all the student's friends were trouble-makers, the student's PISA reading score was 36 points lower on average, and the probability of an overall average grade of at least 70 percent was 0.16 lower (relative to a base probability of 0.58). Interestingly, if most, but not all, of one's friends were perceived as trouble-makers, the negative effect on academic performance was often larger.

The results also suggest that if the student attended a school where the principal considered bullying to be a problem, the overall average grade, as well as the math and language grades, were significantly lower. Taking math grades as an example, the probability of obtaining a grade of 70 percent or better fell by 0.07 if bullying was a problem, which constituted a 13 percent decrease relative to the base case probability of 0.53. However, the existence of a bullying problem had no impact whatsoever on PISA scores. Similarly, the lack of significant coefficients for the drug and alcohol use variable (*drugs*) and the poor home environment variable (*poorhome*) in any of the seven models seemed to contradict the principal's view that these factors were inhibiting learning at the school.

Overall, the results reported so far confirm the correlation between academic performance and individual, family background, and peer/neighbourhood characteristics. But what about school resources, the key variables for this study? Considering Table 8 first, the results show that very few of the remaining school characteristics included in the models had a statistically significant coefficient. The indicator for a co-educational school (*coed*) was the only such variable that had a statistically significant coefficient in the overall average grade model; it was positively related to overall average grades, but only at the 10 percent level of significance. This coefficient implies that if the student attended a co-educational school, holding all else equal, the probability that the student would have received a grade above 69 percent would have increased by 0.09. Furthermore, the results in Table 9 show that attending a co-educational school did not have a significant impact on PISA scores, which were arguably a better measure of student proficiency.

The student-teacher ratio (*stratio*), a popular measure of school resources in other studies, had a negative and significant (at the 5 percent level) coefficient in just one model, that for language grades. Its value implied that if the number of students per teacher were increased by one, the probability of their receiving a grade above 69 percent would have increased by only 0.006 (relative to a base probability of 0.48), a small effect of less than one percentage point.

The only other school characteristics that had a significant impact on reported grades at the 10 percent or a lower level of significance were school size (*schlsize*), which had a negative impact on language grades, and the proportion of science teachers with a bachelor's degree in science (*propscie*), which had an unexpected negative impact on science grades. According to the estimates, an increase of 0.1 in the proportion of science teachers with a bachelor's degree in science was associated with an increase of 0.007 in the probability of obtaining a grade of at least 70 percent.

Interestingly, the set of school characteristics that was associated with PISA scores was quite different from that associated with grades. The results suggest that school enrolment (*schlsize*) and school type had the most important effects on PISA scores. All three PISA scores were significantly influenced by school size and whether or not the school had a religious affiliation. Attending a private school was associated with an increase of 12 points in the PISA math score at the 10 percent level of significance. However, there was no evidence of a significant difference between public and private school students in the PISA science or reading scores[19]. Attending a school with a religious affiliation was associated with a corresponding decrease in PISA scores of about 13 to 16 points, depending on the domain. In contrast, school size actually had a positive, not negative, effect on student performance, although the effect was not large in magnitude; an increase in enrolment of 100 students was associated with an increase in PISA scores of less than 2 points.

Finally, one other school-related variable did seem to have some effect on PISA reading scores at the 5 percent level of significance. The higher the index of teacher participation, the greater the PISA reading score. However, the coefficient estimate implies that a 10 percent increase in the index from its Canadian average value of 0.71 would have been associated with an increase of only 0.19 in

---

19 In preliminary work, attending a private school was also significantly related to PISA scores in science and reading. However, the addition of *avefaminc*, *puni*, and *ppemp* to the model greatly reduced the private school effect.

the PISA reading score, an effect which could not be considered large[20].

Overall, these results with respect to the effect of school resources on student performance were similar to those obtained elsewhere, in that they revealed little or no detectable effect of school resources on academic performance. The most consistent results were those for school size and school type (religious affiliation), which were consistent across the three PISA score models. It is possible that the failure to observe effects of other measures of school resources may have been due to a lack of variation in these measures across schools in the sample.

The Effect of High School Resources on Educational Attainment

The results outlined in the previous section showed that very few school characteristics had an impact on PISA scores or high school grades at age 15. This section of the paper tests whether the characteristics of the school attended at age 15 had any effect on educational attainment at ages 19 and 21. As noted above, at each age, two alternative categorical variables measuring educational attainment were available, one of which focused on the individual's attainment with respect to high school graduation (HEDLD3 and HEDLD4) and the other focusing on the type of post-secondary program attempted (ATTAIN3 and ATTAIN4). In all four models, the PISA reading score (*wleread*) was included as a measure of ability[21].

The coefficient estimates for these four ordered probit models are presented in Table 10 and Table 12. Note that the reference individual used in these models is the same as that used in the models estimated in the previous section. It should also be noted that, due to attrition over time, the sample sizes for these models were somewhat smaller than those in Table 8.

The first thing to note about the estimates shown in Table 10 is that the results with respect to individual characteristics were fairly similar to those shown in Table 8 and Table 9. This finding may reflect the fact that the student's opportunities at each subsequent stage of educational attainment were determined by his or her grades. Although it is possible that a student's grades might change over time, the similarity in the results also suggests that, for most students, grades at age 15 were a good predictor of grades later in life.

The one interesting exception with respect to personal characteristics was that, although learning disabilities were strongly negatively related to both grades and PISA scores at age 15, having a learning disability seemed to have less of an effect on later educational attainment. The learning disabilities indicator (*lrndis*) did not have a significant coefficient in any of the four models of educational attainment. Instead, "other" disabilities seemed to play a more important role in influencing educational attainment beyond the high school level, with their ef-

20 The Canadian mean value was obtained from p. 246 of OECD (2003).

21 Because the PISA math and science tests were not administered to all survey participants, the sample size would have decreased considerably if all three PISA scores were included. When one model was estimated with both the PISA math and reading scores included, the sample size dropped by more than 50%.

fects being negative and statistically significant in all four models of educational attainment. Having a disability in this category reduced the probability of attending post-secondary education by 0.08 to 0.11, depending on the model, with the effect being smallest for HEDLD3 and largest for ATTAIN4. Having a physical disability, on the other hand, significantly increased the probability of having taken some post-secondary education by age 21, but only in the model with HEDLD4 as the dependent variable[22].

With respect to the family background variables, many of the results – particularly those with respect to the parent's educational aspirations for the student, the number of times the child had changed homes prior to age 15 (*nmoves*), and parental education – were again very similar. The higher the level of education attained by a student's parents, the higher his or her level of educational attainment was likely to be later in life. Surprisingly, the HEDLD4 model showed a less pronounced impact of parental education on educational attainment, even though parental aspirations were still important. This finding suggests that high schools may have done a good job of offsetting inequalities in student aspirations due to differences in parental education levels. However, the strong effects of parental education in the model for ATTAIN4 suggest that parental education still had an important effect on the level of PSE attained. Finally, the strong effect of birth order that was observed at age 15 no longer seemed to be present at ages 19 and 21.

Once again, there were also some provincial effects, although these were not necessarily consistent with those for the academic performance equations. The coefficient of the Alberta dummy variable (*alta*) was negative and significant in the models for educational attainment as of age 19, implying that young people raised in Alberta were less likely to pursue post-secondary education than those in Ontario. This finding seemed somewhat surprising given Alberta students' strong levels of proficiency (as measured by the PISA), but may have reflected the strong labour market in Alberta over the past decade. The booming oil industry in that province has raised wages even for unskilled jobs, making post-secondary education a less attractive proposition. By age 21, however, the coefficient of the Alberta dummy variable was no longer statistically significant, suggesting that, only two years later, young people in Alberta had caught up to their counterparts in Ontario. On the other hand, the coefficient of the dummy variable for Quebec had a statistically significant coefficient only in the models that did not distinguish between different types of PSE. When these simpler measures of attainment were used, the coefficient was positive.

Another interesting finding was that peer effects, as measured by the proportion of one's friends that were considered to be trouble-makers, and the three school climate variables (*bullying*, *drugs*, and *poorhome*), seemed to dissipate over time. While some of these variables still had some statistically significant effects by age 21, there were fewer cases in which this was true than earlier on, at age 15 or age 19. In fact, the three

22 In this model the effect is actually fairly large, amounting to 0.19, as compared to a base probability of 0.39.

school climate variables had no effect at all on educational attainment at ages 19 and 21.

However, the variables of primary interest in these models were the school characteristics and the PISA reading score. The PISA reading score (*wleread*) had a positive coefficient that was significant at the 1 percent level in all equations, implying that strong cognitive skills, as measured by this variable, increased the probability of attending post-secondary education. Focusing on the third equation in Table 10, the coefficient estimate implies that an increase of 100 points in an individual's PISA reading score at age 15 would have increased the probability of his or her attempting a post-secondary program by age 19 by 0.15 (with a base probability of 0.34). The fourth model implies that the probability of his or her attempting a bachelor's degree program would have risen by 0.07 (with a base probability of 0.09), constituting an increase of 78 percent.

The significance of the coefficient of the PISA reading score implies that, even if school characteristics did not have a significant direct impact on educational attainment, any characteristic that influenced the PISA reading score in Table 9, such as school size, would have had an indirect impact on educational attainment. In addition, looking again at Table 10, several school characteristics appear to have had an independent effect on at least one measure of educational attainment. Their marginal effects are summarized in Table 11.

Attending a private school was associated not only with higher PISA math scores, but also with higher levels of educational attainment at ages 19 and 21. However, that this effect was observed only in the models that did not distinguish between the different levels of PSE suggests that attending a private school had more of an influence on whether or not the student graduated from high school than on the level of PSE taken. The results imply that the probability of attempting some form of post-secondary education was 0.18 higher at age 19 and 0.19 higher at age 21 for students who had attended private schools at age 15.

In addition to having attended a private school, two other school variables were found to have had a statistically significant effect in both of the two simpler models of educational attainment: the proportion of teachers with a bachelor's degree in education (*propqual*) and the index of the quality of the school's educational resources (*scmatedu*). The marginal effects in Table 11 indicate that increases in the proportion of qualified teachers were associated with decreases in the proportion of individuals that attempted post-secondary education. Although the marginal effects may seem large here, it should be remembered that the associated variable ranged from 0 to 1, and its average value was over 0.9. Hence a reasonable change in this variable would probably have been about 0.01. If the marginal effects in the table were multiplied by this value, they would indeed have been much smaller, so the unexpected negative sign of the effect may be less important than it seems. The marginal effect of a change in the index of school educational resources was also small.

In the models with 11 levels of educational attainment, even fewer measures of school resources seem to have had a statistically significant effect. In fact, in the model of educational attainment at age 19 with ATTAIN3 as the dependent variable, none of the school resource variables had significant coefficients. It was therefore somewhat surprising that three such variables did have significant coefficients at age 21, when ATTAIN4 was the dependent variable. Of these, attending a religious school appears to have had the largest marginal effect; the results in the fourth column of Table 10 suggest that attending a religious school was associated with a higher probability of having taken some PSE by age 21, with the predicted increase in the probability of having done so being 0.06. The lion's share of this increase occurred in just two categories of PSE, college or CEGEP programs and bachelor's degree programs. In both cases, the probability of having taken the program was 3 percentage points higher for individuals who attended religious schools at age 15. This result was unexpected, since attendance at a religious school was associated with lower, not higher, PISA scores and appeared to be unrelated to high school grades at age 15.

Overall, the results with respect to school resources were somewhat mixed. Only one of the school resource variables had a statistically significant coefficient in more than two models. Furthermore, the effects that were observed were sometimes counterintuitive (i.e., the sign of the coefficient was the opposite of that which had been predicted) and, with the exception of the school type effects, seemed small. Coupled with the fact that the school measures represented only one point early in the student's high school career, it would be difficult to say with any confidence that the observed effects were truly causal in nature.

Conclusions

The YITS-A is an extremely rich data set without which it would have been impossible to make this contribution to the extremely limited literature on the effects of school resources on student outcomes in Canada. It contains many possible outcome measures and many possible measures of school resources, only a small subset of which were included in the models estimated in this paper. Nonetheless, it does suffer from several limitations, a notable one being that information on school characteristics was only collected at one point in time – for the year in which the student turned 15. Information on family characteristics was also only made available for that year. Another limitation is that only very limited information about the neighbourhood in which the individual lived and in which the student's school was located was made available. This limitation made it difficult to control completely for nonrandomness in the allocation of students to schools, leaving open the possibility that the parameter estimates may have been somewhat biased.

Nonetheless, there are so few published studies of the effect of school characteristics on student outcomes in Canada that this study can still be regarded as a useful contribution. Indeed, the present analysis has generated a number of interesting results. First of all, parental education levels and parental aspirations for their children were found to be highly correlated with student performance at age 15 and educational attainment in later years. Birth order was also found to have a relatively large impact on student performance at age 15. Receiving more help with school work from parents and siblings was not actually associated with more positive outcomes for the student, however, perhaps because by age 15 only weaker students would need much extra help.

As far as school resources were concerned, the results of this study were as inconsistent as those in the literature for other countries (Hanushek 2006). Most of the included measures of school resources did not have a significant impact on either reported grades or PISA scores. However, PISA scores at age 15 did seem to be strongly related to certain school characteristics, in particular whether the school had a religious affiliation and school size.

Attending a school with a religious affiliation at age 15 tended to reduce PISA scores and may also have had an independent effect on educational attainment at ages 19 and 21, depending on how educational attainment was measured. Attending a larger school, however, tended to increase PISA scores, albeit slightly, an effect that seemed inconsistent with popular opinion among the general population. It is possible that schools with a large population of students also tended to have more educational resources of various sorts at their disposal. Finally, attending a private school at age 15 had a relatively large positive effect on the probability of attempting post-secondary education in two of the four models of educational attainment, although it didn't seem to have much effect on academic performance at age 15.

These findings raise an interesting question that this study is unable to answer: why should the type of school attended by the student have such an important impact on his or her performance and educational attainment? It seems unlikely that differences in the average quality of teachers or resources in different types of schools can completely answer this question, since other variables reflecting teacher quality, student-teacher ratios, and the quality of the school's physical and educational resources were also included in the present analysis.

There exist at least two possible alternative explanations. The first is that there existed some substantive differences between the students who attended the different types of schools that were not captured by the family and student characteristics already included in the model. Exactly what these differences might have been is not clear, since a number of variables that might have been expected to be correlated with attendance at private schools, such as family income and parental education, were already included. Another possible explanation is that private schools, public schools, and schools with a religious affiliation may somehow have used their available resources differently. In other words, their education production functions may have been different.

This latter explanation fits well with Hanushek's (1996, 69) assertion that "the central issue in all policy discussions is usually not whether to spend more or less on school resources but how to get the most out of marginal expenditures." Hanushek (1996, 2006) went on to argue that more attention should be devoted to the incentives facing both students and teachers. Could it be that private schools somehow offer students stronger incentives to perform well? Do schools with a religious affiliation do the reverse? These are questions worth investigating, but other types of data may be needed in order to do so.

To return to the policy question raised in the introduction, unfortunately the results herein did not provide a clear answer to questions about the return to investing more financial resources at the secondary school level. At most, the results supported the concentration of students in larger schools, since PISA scores seemed to be positively related to school size. However, as noted earlier, the effect of school size was very small. It would be difficult to draw any policy conclusions from the observed positive benefits of attending private schools, since the mechanisms that gave rise to these benefits is unknown. The observed benefits do not seem to have stemmed from greater school autonomy or better teacher morale, since controls for these factors were included in the models.

One possible explanation for the failure of this study to detect strong effects of traditional measures of school resources on educational performance may simply be that there was insufficient variation in these measures across schools in Canada. The small differences between the averages of these variables across levels of performance that are evident in Tables 3 and 4 certainly suggest that this may be the case.

Another possibility is that the models did not include enough information about school resources. The cumulative effect of the student's experiences in school prior to age 15 may have had a much greater impact on academic performance at age 15 than the resources available at the school attended in that particular year. While the National Longitudinal Survey of Children and Youth did provide some information about school characteristics at earlier ages, it too may not have collected enough information to properly address this issue[23].

Finally, as Corak and Lauzon (2005) suggested, the effects of school resources may differ across the population. However, estimation results for different income classes (not presented here), as well as for subsets of students with high and low PISA scores, did not yield very different results. Nonetheless, further work along these lines might yield additional insights, as might further investigation of the sources of the observed school type effects.

The results with respect to school resources did not lead to any clear policy recommendations regarding which resources had the greatest impact; however, some of the other results did suggest alternative strategies that policymakers could use to improve student performance in high school as well as educational attainment later in life. First, there was a clear negative relationship between all the educational outcome measures examined and the frequency of smoking. Whether or not this relationship was causal is not clear but, if so, an effective anti-smoking campaign aimed at teenagers might result in

23 The NLSCY also includes far fewer observations than the YITS-A.

improved educational as well as health outcomes. Similarly, the results suggest that any policy that succeeded in raising parents' educational aspirations for their children would also improve educational outcomes in both the short run and the long run. In addition, the results suggest that early interventions designed to discourage teens from becoming "troublemakers" may benefit not only the targeted students, but also their peers, since the results indicate that students with a low proportion of "troublemakers" among their friends have better educational outcomes. Finally, improvements in methods of treating mental health disorders, which were included in the "other disabilities" category, might also result in better educational outcomes in the long run for those who suffer from such disabilities. While neither the design nor the implementation of such policies would be easy, the results of this study suggest that improvements in these areas might have a greater impact on educational outcomes in Canada than marginal improvements in high school resources.

References

Akerhielm, K. 1995. "Does Class Size Matter?" Economics of Education Review 14 (3): 229-241.

Bedard, K. 2003. "School Quality and the Distribution of Male Earnings in Canada." Economics of Education Review 22: 395-407.

Betts, J.R. 1996. "Is There a Link between School Inputs and Earnings? Fresh Scrutiny of an Old Literature," in Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success, ed. G. Burtless. Washington, D.C.: Brookings Institution, pp. 141-191.

Betts, J.R., and D. Morell. 1999. "The Determinants of Undergraduate Grade Point Average." Journal of Human Resources, 24 (2): 268-293.

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

Card, D., and A.B. Krueger. 1996. "Labour Market Effects of School Quality: Theory and Evidence," in Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success, ed. G. Burtless. Washington, D.C.: Brookings Institution, pp. 97-140.

Clotfelter, C.T., H.F. Ladd, and J.L. Vigdor. 2007. "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." NBER Working Paper 13617, National Bureau of Economic Research, Cambridge, MA, November.

Corak, M., and D. Lauzon. 2002. "Provincial Differences in High School Achievement: For Whom do Schools Matter?" Proceedings of Statistics Canada Symposium 2002: Modeling Survey Data for Social and Economic Research. At http://www.statcan.ca/english/freepub/11-522-XIE/2002001/session9/lauzon.pdf.

Corak, M., and D. Lauzon. 2005. Differences in the Distribution of High School Achievement: The Role of Class Size and Time-in-term. Analytical Studies Branch Research Paper Series, Statistics Canada, Catalogue no. 11F0019MIE – No. 270, November.

Dearden, L., J. Ferri, and C. Meghir. 2002. "The Effect of School Quality on Educational Attainment and Wages." Review of Economics and Statistics, 84 (1): 1-20.

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

Eide, E.R., and M.H. Showalter (2005) "Does Improving School Quality Reduce the Probability of Unemployment?" Contemporary Economic Policy 23 (4), 578-584.

Finnie, R., and R. Meng. 2002. "Minorities, Cognitive Skills and Incomes of Canadians." Canadian Public Policy 28 (2), 257-273.

Frenette, M. 2007. "Why Are Youth from Lower-income Families Less Likely to Attend University? Evidence from Academic Abilities, Parental Influences, and Financial Constraints." Research Paper No. 295 (cat. no. 11F0019MIE), Analytical Studies Branch, Statistics Canada.

Green, D.A., and W. C. Riddell. 2003. “Literacy and Earnings: An Investigation of the Interaction of Cognitive and Unobserved Skills in Earnings Generation.” Labour Economics 10, 165-184.

Hanushek, E.A. 1996. “School Resources and Student Performance,” in Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success, ed. G. Burtless. Washington, D.C.: Brookings Institution, pp. 43-73.

Hanushek, E.A. 2006a. “School Resources,” in Handbook of the Economics of Education, Vol. 2, eds. E.A. Hanushek and F. Welch. Amsterdam: Elsevier, pp. 865-908.

Hanushek, E.A., J.F. Kain, J.M. Markman, and S.G. Rivkin. 2003. “Does Peer Ability Affect Student Achievement?” Journal of Applied Econometrics 18, 527-544.

Hanushek, E.A., and S.G. Rivkin. 2006. “Teacher Quality,” in Handbook of the Economics of Education, Vol. 2, eds. E.A. Hanushek and F. Welch. Amsterdam: Elsevier, pp. 1051-1078.

Heckman, J., A. Layne-Farrar, and P. Todd. 1996. “Does Measured School Quality Really Matter? An Examination of the Earnings-Quality Relationship,” in Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success, ed. G. Burtless. Washington, D.C.: Brookings Institution, pp. 192-289.

Hedges, L.V., and R. Greenwald. 1996. “Have Times Changed? The Relation between School Resources and Student Performance,” in Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success, ed. G. Burtless. Washington, D.C.: Brookings Institution, pp. 74-92.

Hedges, L.V., R.D. Laine, and R. Greenwald. 1994. “Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes.” Educational Researcher 23 (April): 5-14.

Heinesen, E., and B.K. Graversen. 2005. “The Effect of School Resources on Educational Attainment: Evidence from Denmark.” Bulletin of Economic Research 57 (2): 109-143.

Johnson, D. 2005. Signposts of Success: Interpreting Ontario’s Elementary School Test Scores. C. D. Howe Institute Policy Study 40. Ottawa: Renouf Publishing.

Ludwig, J., H.F. Ladd, and G.J. Duncan. 2001. “Urban Poverty and Educational Outcomes.” Brookings-Wharton Papers on Urban Affairs 2001: 147-201.

Marcotte, D.E. 2007. “Schooling and Test Scores: A Mother-Natural Experiment.” Economics of Education Review 26: 629-640.

Murnane, R.J., J.B. Willett, M.J. Braatz, and Y. Duhaldeborde. 2001. “Do Different Dimensions of Male High School Students’ Skills Predict Labor Market Success a Decade Later? Evidence from the NLSY.” Economics of Education Review 20: 311-320.

Nechyba, T.J. 2006. “Income and Peer Quality Sorting in Public and Private Schools,” in Handbook of the Economics of Education, Vol. 2, eds. E.A. Hanushek and F. Welch. Amsterdam: Elsevier, pp. 1327-1368.

OECD. 2002. Manual for the PISA 2000 Database. Centre for Research and Innovation, SourceOECD Education, Vol. 2002, No.4.

OECD. 2003. PISA 2000 Technical Report. Centre for Research and Innovation, SourceOECD Education, Vol. 2002, No.11.

Rangvid, B.S. 2007. "School Composition Effects in Denmark: Quantile Regression Evidence from PISA 2000." Empirical Economics 33: 359-388.

Rivkin, S.G., E.A. Hanushek, and J.F. Kain. 2005 "Teachers, Schools, and Academic Achievement." Econometrica 73 (2): 417-458.

Solon, G., M.E. Page, and G.J. Duncan. 2000. "Correlations between Neighboring Children in Their Subsequent Educational Attainment." Review of Economics and Statistics 82 (3): 383-392.

Speakman, R., and F. Welch. 2006. "Using Wages to Infer School Quality," \ in Handbook of the Economics of Education, Vol. 2, eds. E.A. Hanushek and F. Welch. Amsterdam: Elsevier, pp. 813-864.

Wilson, K. 2001. "The Determinants of Educational Attainment: Modeling and Estimating the Human Capital Model and Education Production Function." Southern Economic Journal, 67(3): 518-551.

Tables and Figures

Table 1. Variable definitions

Variable Name	Definition
fem	Equal to 1 if individual is female, 0 otherwise
vismin	Equal to 1 if individual belongs to visible minority group, 0 otherwise
imm	Equal to 1 if individual is an immigrant, 0 otherwise
ysm1	Years since immigration for immigrants, 0 for Canadian born (as of cycle one)
ysm3	Years since immigration for immigrants, 0 for Canadian born (as of cycle three)
ysm4	Years since immigration for immigrants, 0 for Canadian born (as of cycle four)
physdis	One if the individual has a physical disability that affects schooling, 0 otherwise
phealthdis	One if the individual has a health problem that affects schooling, 0 otherwise
lrndis	One if the individual has a learning disability that affects schooling, 0 otherwise
othdis2	One if the individual has another type of disability (including a mental health condition) that affects schooling, 0 otherwise
smoker	One if the individual smokes at least once per week, 0 otherwise
occsmoker	One if the individual is a smoker but smokes less than once per week, 0 otherwise
nschl1	One if number of schools attended by age 15 is 1, 0 otherwise
nschl2	One if number of schools attended by age 15 is 2, 0 otherwise
nschl4	One if number of schools attended by age 15 is 4, 0 otherwise
nschl5	One if number of schools attended by age 15 is 5, 0 otherwise
nschl6	One if number of schools attended by age 15 is 6 or more, 0 otherwise
nsib0	Equal to 1 if no siblings, 0 otherwise
nsib2	Equal to 1 if 2 siblings, 0 otherwise
nsib3	Equal to 1 if 3 siblings, 0 otherwise
nsib4	Equal to 1 if 4 or more siblings, 0 otherwise
oldsib1	Equal to 1 if 1 older sibling, 0 otherwise
oldsib2	Equal to 1 if 2 older siblings, 0 otherwise
oldsib3	Equal to 1 if 3 older siblings, 0 otherwise
oldsib4	Equal to 1 if 4 older siblings, 0 otherwise
nmoves	Number of times child had changed homes by age 15
faminc	Family income in 1999 (dollars)
famedsup	Index of family educational support
twop	Equal to 1 if lived in two-parent household at age 15, at most one biological parent; 0 otherwise
singlep	Equal to 1 if lived in single-parent household at age 15, 0 otherwise
nohs	Equal to 1 if neither parent completed high school, 0 otherwise
somepse	Equal to 1 if at least one parent attended college or university, but did not receive degree or diploma; 0 otherwise
privpse	Equal to 1 if at least one parent has a diploma from a private post-secondary institute, 0 otherwise
diploma	Equal to 1 if at least one parent has a college diploma or some other non-university credential, 0 otherwise
ucert	Equal to 1 if at least one parent has a university certificate, 0 otherwise
bach	Equal to 1 if at least one parent has a bachelor's degree, 0 otherwise
prof	Equal to 1 if at least one parent has a professional degree, 0 otherwise
grad	Equal to 1 if at least one parent has a graduate degree, 0 otherwise
etrade	One if parent hopes child will achieve at least a trade qualification, 0 otherwise
ecoll	One if parent hopes child will attend college, 0 otherwise

Table 1 continued

Variable Name	Definition
euni	One if parent hopes child will attend university, 0 otherwise
epse	One if parent hopes child will attend some form of PSE, 0 otherwise
sometrbl	One if some of student's friends are trouble-makers, 0 otherwise
mosttrbl	One if most of student's friends are trouble-makers, 0 otherwise
alltrbl	One if all of student's friends are trouble-makers, 0 otherwise
nfld	Equal to 1 if province of schooling is Newfoundland and Labrador, 0 otherwise
pei	Equal to 1 if province of schooling is Prince Edward Island, 0 otherwise
ns	Equal to 1 if province of schooling is New Scotia, 0 otherwise
nb	Equal to 1 if province of schooling is New Brunswick, 0 otherwise
que	Equal to 1 if province of schooling is Quebec, 0 otherwise
man	Equal to 1 if province of schooling is Manitoba, 0 otherwise
sask	Equal to 1 if province of schooling is Saskatchewan, 0 otherwise
alta	Equal to 1 if province of schooling is Alberta, 0 otherwise
bc	Equal to 1 if province of schooling is British Columbia, 0 otherwise
rural	Equal to 1 if student lived in village or rural area at age 15, 0 otherwise
stown	Equal to 1 if student lived in small town at age 15, 0 otherwise
town	Equal to 1 if student lived in a town at age 15, 0 otherwise
ccentre	Equal to 1 if student lived close to centre of large city at age 15, 0 otherwise
ncentre	Equal to 1 if student lived in a large city at age 15, but not near city centre; 0 otherwise
schlsize	Total enrolment in school
Coed	Equal to 1 if school is co-educational, 0 otherwise
Private	Equal to 1 if school is private, 0 otherwise
relschl	Equal to 1 if school has religious affiliation, 0 otherwise
Stratio	Student-teacher ratio of school
propqual	proportion of teachers with a degree in education
propmath	proportion of math teachers with a bachelor's degree in math
propscie	proportion of science teachers with a bachelor's degree in science
propread	proportion of language teachers with a bachelor's degree in that subject
scmatbui	index of quality of school physical resources
scmatedu	index of quality of school educational resources
skills	Equal to 1 if school offers special courses in study skills, 0 otherwise
ratcomp	Ratio of computers to students in school
bullying	One if bullying interferes with learning at school, 0 otherwise
drugs	One if drug and alcohol use interfere with learning at school, 0 otherwise
poorhome	One if poor home environments interfere with learning at school, 0 otherwise
numhs_c3	Number of high schools attended as of cycle three
numhs_c4	Number of high schools attended as of cycle four
wleread	PISA reading score

Table 2. Students by school type

	Private Schools	Public Schools	All Schools
Religious Affiliation	12857	56260	69116
No Religious Affiliation	2585	127065	195865
Total	15441	183325	198767

Source: Author's calculations using data from the YITS-A.
The figures are population estimates constructed using data for one of the samples used to for estimation and sampling weights.

Table 3. Students by overall average grade and school type

Overall Average Grade at Age 15	Private School	Public School	All Schools
59% or less	384	16040	16424
60%-69%	2049	33694	35744
70%-79%	5252	57062	62314
80%-89%	5379	51031	56410
90%-100%	1320	12670	13991
Total	14385	170497	184882

Source: Author's calculations using data from the YITS-A.
The figures are population estimates constructed using one of the samples used for estimation and sampling weights.
The total differs from that in Table 2 because not all students for whom the PISA reading score was available reported their overall average grade.

Table 4. Students by PISA reading score and school type

Overall Average Grade at Age 15	Private School	Public School	All Schools
59% or less	384	16040	16424
60%-69%	2049	33694	35744
70%-79%	5252	57062	62314
80%-89%	5379	51031	56410
90%-100%	1320	12670	13991
Total	14385	170497	184882

Source: Author's calculations using data from the YITS-A.
The means are population estimates constructed using data for the sample used to estimate the last equation in Table 9 and sampling weights.

Table 5. Average school resources by overall average grade at age 15

Variable	< 50%	50%-59%	60%-69%	70%-79%	80%-89%	90%-100%
schlsize	931	964	1000	987	1003	945
stratio	17.0	17.2	17.2	17.1	17.0	17.0
propqual	0.96	0.94	0.93	0.92	0.92	0.95
scmatbui	-0.06	-0.31	-0.34	-0.36	-0.34	-0.27
scmatedu	0.01	-0.15	-0.23	-0.23	-0.22	-0.13
ratcomp	0.18	0.19	0.18	0.19	0.19	0.19
skills	0.70	0.71	0.68	0.69	0.67	0.68

Source: Author's calculations using data from the YITS-A.
The means are population estimates constructed using data for the sample used to estimate the last equation in Table 9, adjusted to exclude those who did not report their overall average grade, and sampling weights.

Table 6. Average school resources by PISA reading level at age 15

Variable	Level 0	Level 1	Level 2	Level 3	Level 4	Level 5
schlsize	687	873	940	1001	1009	1041
stratio	15.1	16.5	16.9	17.1	17.1	17.4
propqual	0.95	0.95	0.94	0.92	0.92	0.92
scmatbui	-0.44	-0.32	-0.35	-0.34	-0.34	-0.32
scmatedu	-0.17	-0.05	-0.14	-0.21	-0.26	-0.27
ratcomp	0.22	0.19	0.19	0.18	0.19	0.19
skills	0.60	0.73	0.68	0.69	0.67	0.66

Source: Author's calculations using data from the YITS-A.
The means are population estimates constructed using data for the sample used to estimate the last equation in Table 9 and sampling weights.

Table 7. Estimated population means of explanatory variables

Variable	Mean
fem	0.50
vismin	0.11
imm	0.07
ysm1	0.50
nschl1	0.12
nschl2	0.28
nschl4	0.16
nschl5	0.07
nschl6	0.05
nsib0	0.07
nsib2	0.30
nsib3	0.14
nsib4	0.11
oldsib1	0.38
oldsib2	0.17
oldsib3	0.06
oldsib4	0.04
nmoves	2.18
faminc	$69200
famedsup	-0.04
twop	0.13
singlep	0.16
nohs	0.10
somepse	0.07
privpse	0.03
diploma	0.28

Variable	Mean
ucert	0.05
bach	0.17
prof	0.02
grad	0.07
nfld	0.02
pei	0.004
ns	0.04
nb	0.02
que	0.24
man	0.04
sask	0.05
alta	0.11
bc	0.10
rural	0.11
stown	0.23
town	0.26
ccentre	0.05
ncentre	0.04

Source: Author's calculations using data from the YITS-A.
The means are population estimates constructed using data for the sample used to estimate the last equation in Table 9 and sampling weights.

Table 8. Ordered probit models of high school grades at age 15

	Dependent Variable			
Variable Name	**Overall Average**	**Math**	**Science**	**Language**
Student characteristics				
fem	0.336***	0.0717***	0.111***	0.511***
	[0.033]	[0.027]	[0.031]	[0.032]
vismin	0.026	0.0293	0.0431	-0.0852
	[0.06]	[0.067]	[0.06]	[-0.055]
imm	-0.0454	0.482***	0.162	-0.521***
	[0.15]	[0.16]	[0.15]	[0.12]
ysm1	0.00833	-0.0432**	-0.0283*	0.0642***
	[0.018]	[0.018]	[0.017]	[0.014]
physdis	0.117	-0.0274	-0.131	-0.115
	[0.14]	[0.14]	[0.12]	[0.13]
phlthdis	-0.0957	0.0539	-0.0633	0.0144
	[0.096]	[0.085]	[0.093]	[0.088]
lrndis	-0.682***	-0.293**	-0.759***	-0.455***
	[0.15]	[0.14]	[0.13]	[0.13]
othdis2	0.0135	-0.0954	0.0276	0.0139
	[0.095]	[0.091]	[0.083]	[0.086]
smoker	-0.768***	-0.429***	-0.563***	-0.474***
	[0.041]	[0.041]	[0.037]	[0.042]
occsmoker	-0.306***	-0.217***	-0.248***	-0.169***
	[0.042]	[0.043]	[0.042]	[0.04]
Number of schools attended				
nschl1	-0.144***	-0.0684	-0.128***	-0.143***
	[0.053]	[0.047]	[0.049]	[0.05]
nschl2	-0.0319	-0.041	-0.0706**	-0.034
	[0.037]	[0.035]	[0.036]	[0.038]
nschl4	0.0558	0.00697	0.00713	0.0175
	[0.047]	[0.044]	[0.045]	[0.043]
nschl5	-0.00327	0.0134	0.0303	0.0227
	[0.067]	[0.062]	[0.062]	[0.06]
nschl6	-0.0453	0.00123	-0.0869	0.0353
	[0.086]	[0.074]	[0.07]	[0.076]

Table 8 continued

	Dependent Variable			
Variable Name	**Overall Average**	**Math**	**Science**	**Language**
Number of siblings				
nsib0	-0.134**	-0.170***	-0.126**	-0.0338
	[0.059]	[0.058]	[0.063]	[0.058]
nsib2	0.0144	0.0137	-0.0670*	0.00916
	[0.038]	[0.037]	[0.035]	[0.04]
nsib3	-0.0229	0.043	-0.0144	0.0289
	[0.052]	[0.044]	[0.043]	[0.046]
nsib4	-0.0194	0.123*	-0.0473	0.00349
	[0.073]	[0.068]	[0.067]	[0.073]
Number of older siblings				
oldsib1	-0.0632*	-0.0366	-0.109***	-0.0742**
	[0.033]	[0.034]	[0.032]	[0.035]
oldsib2	-0.240***	-0.136***	-0.172***	-0.173***
	[0.05]	[0.044]	[0.044]	[0.043]
oldsib3	-0.149*	-0.186**	-0.161**	-0.124*
	[0.078]	[0.073]	[0.07]	[0.068]
oldsib4	-0.204*	-0.259**	-0.309***	-0.293***
	[0.11]	[0.1]	[0.1]	[0.1]
Other family background variables				
nmoves	-0.0255***	-0.0210***	-0.0209***	-0.0296***
	[0.0065]	[0.0066]	[0.0058]	[0.0062]
faminc	2.15E-07	3.15E-08	3.95E-07	2.29E-07
	[3.4E-07]	[3.5E-07]	[2.6E-07]	[2.4E-07]
famedsup	-0.0412**	-0.0447***	-0.0567***	-0.0157
	[0.017]	[0.016]	[0.014]	[0.014]
twop	-0.193***	-0.0574	-0.142***	-0.157***
	[0.052]	[0.043]	[0.044]	[0.043]
singlep	-0.0404	0.0297	-0.0920**	-0.0569
	[0.045]	[0.041]	[0.045]	[0.045]

Table 8 continued

	Dependent Variable			
Variable Name	**Overall Average**	**Math**	**Science**	**Language**
Parental education variables				
nohs	-0.107*	-0.0769	-0.034	-0.131**
	[-0.061]	[-0.058]	[-0.06]	[-0.058]
somepse	0.124*	-0.0555	0.0352	0.0535
	[0.066]	[-0.065]	[0.064]	[0.066]
privpse	0.0923	0.0214	-0.0285	-0.023
	[0.079]	[0.073]	[-0.071]	[-0.067]
diploma	0.0645	0.0169	0.0462	0.0973**
	[0.042]	[0.037]	[0.037]	[0.041]
ucert	0.0959	0.131*	0.114	0.212***
	[0.08]	[0.074]	[0.07]	[0.066]
bach	0.401***	0.270***	0.268***	0.396***
	[0.046]	[0.048]	[0.044]	[0.045]
prof	0.548***	0.328***	0.393***	0.563***
	[0.11]	[0.11]	[0.11]	[0.12]
grad	0.612***	0.414***	0.434***	0.583***
	[0.075]	[0.067]	[0.067]	[0.064]
Parent's educational aspirations for child				
etrade	0.0157	0.012	-0.0101	0.0112
	[0.088]	[0.077]	[0.094]	[0.096]
ecoll	0.295***	0.11	0.145	0.213**
	[0.082]	[0.083]	[0.089]	[0.086]
euni	0.815***	0.523***	0.629***	0.697***
	[0.083]	[0.082]	[0.091]	[0.084]
epse	0.492***	0.270***	0.338***	0.433***
	[0.094]	[0.088]	[0.094]	[0.091]
Trouble-makers among peers				
sometrbl	-0.170***	-0.112***	-0.182***	-0.121***
	[0.029]	[0.03]	[0.029]	[0.029]
mosttrbl	-0.542***	-0.295***	-0.400***	-0.382***
	[0.054]	[0.055]	[0.056]	[0.056]
alltrbl	-0.420***	-0.330***	-0.367***	-0.311***
	[0.1]	[0.11]	[0.091]	[0.1]

Table 8 continued

	Dependent Variable			
Variable Name	**Overall Average**	**Math**	**Science**	**Language**
Province of schooling				
nfld	0.0225	-0.0137	-0.0498	0.128
	[0.098]	[0.088]	[0.084]	[0.1]
pei	0.251***	0.293**	0.111	0.374**
	[0.095]	[0.14]	[0.1]	[0.15]
ns	0.170**	0.109	0.0362	0.268***
	[0.082]	[0.081]	[0.077]	[0.08]
nb	0.214**	0.313***	0.208**	0.302***
	[0.092]	[0.083]	[0.085]	[0.11]
que	0.0632	0.150**	0.256***	0.0851
	[0.078]	[0.073]	[0.071]	[0.084]
man	-0.126	-0.0153	0.00906	0.0646
	[0.081]	[0.072]	[0.07]	[0.065]
sask	0.239***	0.271***	0.230***	0.338***
	[0.08]	[0.075]	[0.07]	[0.076]
alta	-0.301***	-0.228***	-0.255***	-0.148**
	[0.064]	[0.064]	[0.061]	[0.063]
bc	0.0441	0.047	0.166***	0.254***
	[0.069]	[0.063]	[0.062]	[0.067]
Size of community				
rural	0.173**	0.0868	0.135**	0.0412
	[0.069]	[0.065]	[0.067]	[0.064]
stown	0.128**	0.109*	0.114**	0.0408
	[0.056]	[0.061]	[0.051]	[0.05]
town	0.0766	0.0224	0.0113	-0.0437
	[0.058]	[0.052]	[0.053]	[0.053]
ccentre	-0.0664	0.0387	-0.0222	0.0629
	[0.089]	[0.099]	[0.087]	[0.11]
ncentre	-0.145*	-0.104	-0.248**	-0.102
	[0.088]	[0.14]	[0.12]	[0.086]
Characteristics of school-mates' families				
avefaminc	2.67E-07	3.3E-07	-1.7E-06	0.000000991
	[0.0000015]	[0.0000016]	[0.0000015]	[0.0000015]
puni	-0.164	-0.232	0.387**	-0.369**
	[0.18]	[0.2]	[0.16]	[0.16]
ppemp	0.392	0.000421	0.345	-0.449
	[0.44]	[0.42]	[0.39]	[0.41]

Table 8 continued

	Dependent Variable			
Variable Name	**Overall Average**	**Math**	**Science**	**Language**
School characteristics				
schlsize	-0.000042	-0.0000437	-0.0000297	-0.0000891**
	[0.000042]	[0.000051]	[0.000042]	[0.000045]
coed	0.237*	0.0841	0.247	0.226
	[0.14]	[0.22]	[0.16]	[0.16]
private	0.0846	0.147	0.0488	0.12
	[0.097]	[0.1]	[0.086]	[0.085]
relschl	0.0533	0.0348	-0.0477	0.0081
	[0.051]	[0.049]	[0.049]	[0.049]
stratio	-0.00689	-0.0109	0.0037	-0.0156**
	[0.0079]	[0.0084]	[0.0057]	[0.0061]
propqual	0.0164			
	[0.093]			
propmath		0.102		
		[0.073]		
propscie			-0.184**	
			[0.082]	
proplang				-0.0873
				[0.08]
scmatbui	-0.0183	-0.01	0.033	0.0117
	[0.031]	[0.029]	[0.027]	[0.028]
scmatedu	0.0364	0.0192	0.00966	0.0163
	[0.024]	[0.023]	[0.022]	[0.023]
skills	-0.00412	-0.0316	-0.00163	0.0445
	[0.041]	[0.041]	[0.038]	[0.035]
ratcomp	-0.0104	-0.111	0.299	-0.0605
	[0.21]	[0.23]	[0.24]	[0.17]
bullying	-0.120*	-0.176***	-0.068	-0.139*
	[0.066]	[0.052]	[0.056]	[0.079]
drugs	0.0209	0.0203	0.00785	-0.0133
	[0.044]	[0.041]	[0.038]	[0.045]
poorhome	-0.00148	-0.0124	0.0402	-0.0323
	[0.039]	[0.04]	[0.04]	[0.037]
tcmorale	0.0295	-0.0128	-0.00609	-0.0208
	[0.021]	[0.019]	[0.022]	[0.021]
schauton	0.00534	0.0081	-0.0079	0.0356
	[0.025]	[0.028]	[0.024]	[0.028]
tchparti	-0.00107	0.0311*	0.0188	0.0134
	0.017	0.017	0.017	0.018

Table 8 continued

	Dependent Variable			
Variable Name	**Overall Average**	**Math**	**Science**	**Language**
Threshold values				
mu1	-1.722***	-1.632***	-1.354***	-2.255***
	[0.41]	[0.47]	[0.37]	[0.42]
mu2	-0.688*	-0.797*	-0.48	-1.339***
	[0.41]	[0.46]	[0.36]	[0.42]
mu3	0.261	-0.185	0.197	-0.604
	[0.41]	[0.46]	[0.36]	[0.42]
mu4	1.338***	0.514	0.998***	0.276
	[0.41]	[0.46]	[0.37]	[0.42]
mu5	2.651***	1.387***	1.981***	1.532***
	[0.41]	[0.46]	[0.37]	[0.42]
No. of obs.	10290	11716	11639	11303
Log pseudo L	-13626.043	-18931.932	-17788.786	-16317.258
Wald χ^2 [p-value]	2310.84 [0.000]	1391.00 [0.000]	1950.49 [0.000]	2731.60 [0.000]
pseudo R-squared	0.1092	0.0464	0.0698	0.0876

Notes: Values in table are maximum likelihood coefficient estimates, with robust standard errors corrected for clustering at the school level in square brackets. Estimation was carried out using sampling weights from the YITS-A parent questionnaire. One asterisk denotes statistical significance at the 10% level, two asterisks denote statistical significance at the 5% level, and three asterisks denote statistical significance at the 1% level.

Table 9. Linear regression models of PISA scores at age 15

	Dependent Variable		
Variable Name	**Math Score**	**Science Score**	**Reading Score**
Student characteristics			
fem	-11.22***	-2.182	26.21***
	[2.46]	[2.78]	[1.95]
vismin	-18.37***	-13.51**	-12.92***
	[4.63]	[5.61]	[4.09]
imm	1.745	-77.30***	-58.93***
	[13.9]	[11.9]	[9.91]
ysm1	-1.555	6.512***	3.910***
	[1.54]	[1.36]	[1.12]
physdis	-2.809	-23.32	-21.20**
	[9.47]	[14.8]	[10.7]
phlthdis	-8.527	-1.448	-0.442
	[6.40]	[8.52]	[5.83]
lrndis	-43.39***	-50.63***	-62.71***
	[8.97]	[14.0]	[8.52]
othdis2	-15.88**	-1.922	-5.006
	[6.44]	[7.59]	[5.04]
smoker	-28.01***	-23.86***	-31.63***
	[3.33]	[3.30]	[2.81]
occsmoker	-5.785	-10.51**	-4.618
	[4.00]	[4.21]	[2.84]
Number of schools attended			
nschl1	-25.35***	-23.99***	-25.90***
	[4.71]	[4.27]	[3.64]
nschl2	-2.93	-8.487**	-3.984*
	[3.56]	[3.60]	[2.35]
nschl4	3.577	4.789	7.177**
	[3.82]	[4.34]	[3.21]
nschl5	4.437	1.55	6.041
	[5.05]	[5.23]	[4.12]
nschl6	0.26	7.952	5.516
	[6.17]	[8.62]	[6.45]
Number of siblings			
nsib0	-8.347*	-8.114	-11.20***
	[4.65]	[5.78]	[4.28]
nsib2	0.565	-8.346**	-0.861
	[3.26]	[3.97]	[2.67]
nsib3	4.788	-1.045	1.487
	[4.74]	[4.48]	[3.56]
nsib4	9.419*	-3.235	0.446
	[5.72]	[6.73]	[4.67]

Table 9 continued

	Dependent Variable		
Variable Name	**Math Score**	**Science Score**	**Reading Score**
Number of older siblings			
oldsib1	-10.95***	-13.44***	-8.751***
	[3.06]	[3.39]	[2.29]
oldsib2	-15.41***	-16.37***	-16.09***
	[4.27]	[5.27]	[3.40]
oldsib3	-19.57***	-17.59***	-18.95***
	[6.75]	[6.39]	[4.83]
oldsib4	-37.74***	-16.63*	-19.50***
	[8.39]	[9.19]	[6.90]
Other family background variables			
nmoves	-1.680***	-1.508***	-1.451**
	[0.61]	[0.58]	[0.57]
faminc	-7.86E-08	0.00000467	0.00000531
	[0.000033]	[0.000029]	[0.000017]
famedsup	-12.08***	-11.58***	-11.42***
	[1.45]	[1.58]	[1.05]
twop	-13.39***	-12.81***	-7.689**
	[3.96]	[4.10]	[3.07]
singlep	-2.809	7.458**	0.676
	[3.89]	[3.77]	[2.83]
Parental education variables			
nohs	-18.49***	-11.90**	-14.81***
	[5.39]	[5.76]	[4.41]
somepse	-0.563	8.987	9.304*
	[5.69]	[5.77]	[4.77]
privpse	3.63	15.06**	4.894
	[5.48]	[6.83]	[4.68]
diploma	12.42***	17.89***	12.86***
	[3.62]	[3.96]	[2.66]
ucert	19.19***	14.68**	16.73***
	[6.85]	[6.58]	[5.51]
bach	21.55***	34.29***	29.99***
	[4.15]	[4.59]	[3.49]
prof	33.08***	41.99***	41.26***
	[11.2]	[11.2]	[9.35]
grad	41.28***	53.81***	51.20***
	[6.11]	[6.49]	[4.80]

Table 9 continued

	Dependent Variable		
Variable Name	**Math Score**	**Science Score**	**Reading Score**
Parent's educational aspirations for child			
etrade	13.41	10.85	12.67*
	[9.65]	[8.41]	[6.79]
ecoll	19.51**	24.41***	30.60***
	[9.81]	[7.62]	[6.57]
euni	52.35***	51.80***	66.16***
	[9.52]	[7.55]	[6.15]
epse	43.75***	31.08***	48.38***
	[10.7]	[8.93]	[6.66]
Trouble-makers among peers			
sometrbl	-2.933	-3.205	-3.815*
	[2.57]	[2.82]	[2.06]
mosttrbl	-18.06***	-33.94***	-34.23***
	[5.02]	[5.60]	[4.19]
alltrbl	-16.58*	-31.82***	-36.20***
	[9.35]	[12.1]	[9.22]
Province of schooling			
nfld	4.414	-1.283	0.261
	[5.17]	[6.27]	[5.54]
pei	-11.6	-25.92***	-21.00***
	[8.36]	[7.51]	[7.23]
ns	-11.06**	-14.12**	-10.66**
	[5.23]	[5.86]	[5.07]
nb	-10.68*	-26.48***	-18.32***
	[6.46]	[5.93]	[5.62]
que	44.28***	31.22***	24.37***
	[5.19]	[6.47]	[5.01]
man	22.41***	5.221	1.209
	[6.65]	[5.15]	[5.53]
sask	14.18**	2.562	8.036*
	[5.51]	[5.33]	[4.48]
alta	28.18***	29.03***	21.70***
	[5.82]	[5.12]	[4.77]
bc	16.88***	9.857**	11.54***
	[4.85]	[4.39]	[4.10]

Table 9 continued

	Dependent Variable		
Variable Name	**Math Score**	**Science Score**	**Reading Score**
Size of community			
rural	12.52**	4.991	11.52**
	[5.92]	[5.36]	[4.62]
stown	7.264	-2.358	5.656
	[4.49]	[4.96]	[4.69]
town	-2.48	-3.656	1.286
	[3.91]	[4.37]	[3.77]
ccentre	7.167	8.411	0.515
	[14.7]	[10.8]	[10.8]
ncentre	-0.75	-2.125	3.874
	[7.41]	[7.21]	[6.55]
Characteristics of schoolmates' families			
avefaminc	0.000242**	-0.0000911	0.00019
	[0.00011]	[0.00014]	[0.00012]
puni	32.49**	9.462	30.29**
	[13.1]	[13.6]	[13.2]
ppemp	33.67	17.29	41.35
	[26.7]	[33.4]	[25.9]
School characteristics			
schlsize	0.0120***	0.0111***	0.0128***
	[0.0036]	[0.0039]	[0.0040]
coed	7.004	-7.783	-8.893
	[12.0]	[9.22]	[9.96]
private	12.00*	8.909	7.368
	[7.11]	[8.13]	[7.58]
relschl	-12.64***	-16.38***	-12.80***
	[4.24]	[4.31]	[3.64]
stratio	-0.0899	0.43	-0.386
	[0.92]	[0.75]	[0.51]
propmath	-0.687		
	[5.56]		
propscie		5.445	
		[7.19]	
proplang			-4.345
			[6.92]
scmatbui	-2.601	-0.196	1.506
	[2.26]	[2.35]	[2.05]
scmatedu	1.765	2.141	-1.424
	[1.63]	[1.93]	[1.66]
skills	-2.245	-2.474	-1.9
	[3.10]	[3.27]	[2.90]

Table 9 continued

	Dependent Variable		
Variable Name	**Math Score**	**Science Score**	**Reading Score**
School characteristics (cont'd)			
ratcomp	23.53	16.88	10.47
	[15.6]	[16.3]	[15.5]
bullying	-5.022	0.454	-1.934
	[5.21]	[4.47]	[4.26]
drugs	-5.359	0.594	-0.999
	[3.93]	[4.24]	[3.89]
poorhome	3.073	-0.38	-0.356
	[3.08]	[3.21]	[2.99]
tcmorale	-1.086	-0.513	-0.213
	[1.62]	[1.70]	[1.47]
schauton	-0.88	2.145	2.53
	[1.96]	[1.97]	[1.77]
tchparti	0.817	0.819	2.705**
	[1.25]	[1.54]	[1.27]
Constant	427.6***	479.0***	429.2***
	[38.1]	[36.2]	[31.1]
No. of obs.	7058	6981	12162
R-squared	0.2586	0.2343	0.3
F-statistic [p-value]	24.69 [0.000]	29.13 [0.000]	44.75 [0.000]

Notes: Values in table are OLS coefficient estimates, with robust standard errors corrected for clustering at the school level in square brackets. Estimation was carried out using sampling weights from the YITS-A parent questionnaire. One asterisk denotes statistical significance at the 10% level, two asterisks denote statistical significance at the 5% level, and three asterisks denote statistical significance at the 1% level.

Table 10. Ordered probit models of educational attainment

	Dependent Variable			
Variable Name	**hedld3 (4 levels)**	**attain3 (11 levels)**	**hedld4 (4 levels)**	**attain4 (11 levels)**
Student characteristics				
fem	0.353***	0.282***	0.349***	0.303***
	[0.035]	[0.032]	[0.036]	[0.035]
vismin	0.172***	0.203***	0.0476	0.342***
	[0.060]	[0.069]	[0.064]	[0.069]
imm	0.128	0.496***	-0.24	0.372
	[0.20]	[0.19]	[0.21]	[0.29]
ysm1	-0.00304	-0.0187	0.00967	-0.0229
	[0.017]	[0.015]	[0.015]	[0.020]
physdis	0.215	0.198	0.478***	0.288
	[0.19]	[0.19]	[0.16]	[0.18]
phlthdis	-0.132	-0.126	-0.198	-0.121
	[0.100]	[0.12]	[0.14]	[0.12]
lrndis	0.0573	-0.133	0.215	-0.0388
	[0.23]	[0.18]	[0.23]	[0.19]
othdis2	-0.225**	-0.206*	-0.305***	-0.287**
	[0.11]	[0.12]	[0.12]	[0.14]
smoker	-0.460***	-0.421***	-0.345***	-0.408***
	[0.050]	[0.051]	[0.063]	[0.055]
occsmoker	-0.141***	-0.130**	-0.102	-0.114*
	[0.052]	[0.052]	[0.063]	[0.063]
Number of schools attended				
nschl1	-0.201***	-0.166**	-0.270***	-0.236***
	[0.063]	[0.065]	[0.072]	[0.065]
nschl2	-0.00479	-0.0068	-0.0596	-0.0921*
	[0.044]	[0.042]	[0.049]	[0.050]
nschl4	0.0378	-0.0543	-0.000894	-0.135**
	[0.053]	[0.051]	[0.060]	[0.061]
nschl5	0.0922	0.00129	0.016	-0.123
	[0.074]	[0.069]	[0.081]	[0.077]
nschl6	0.0249	-0.0913	0.027	-0.17
	[0.10]	[0.093]	[0.099]	[0.11]
PISA reading score				
wleread	0.00331***	0.00450***	0.00148***	0.00414***
	[0.00026]	[0.00026]	[0.00028]	[0.00030]
Number of high schools attended				
numhs_c3	-0.373***	-0.251***		
	[0.030]	[0.031]		
numhs_c4			-0.322***	-0.261***
			[0.030]	[0.035]

Table 10 continued

	Dependent Variable			
Variable Name	hedld3 (4 levels)	attain3 (11 levels)	hedld4 (4 levels)	attain4 (11 levels)
Number of siblings				
nsib0	-0.015	-0.0123	0.016	-0.11
	[0.089]	[0.076]	[0.091]	[0.077]
nsib2	-0.0217	-0.0728	-0.0184	-0.063
	[0.046]	[0.048]	[0.052]	[0.047]
nsib3	-0.162**	-0.152**	-0.114*	-0.123*
	[0.065]	[0.063]	[0.069]	[0.068]
nsib4	-0.0296	-0.103	-0.0396	-0.123
	[0.081]	[0.077]	[0.098]	[0.097]
Number of older siblings				
oldsib1	0.0266	0.0146	-0.0057	-0.00852
	[0.040]	[0.040]	[0.043]	[0.042]
oldsib2	-0.0579	-0.0173	-0.0237	-0.0646
	[0.058]	[0.056]	[0.061]	[0.061]
oldsib3	-0.102	-0.0961	-0.14	-0.121
	[0.11]	[0.11]	[0.13]	[0.11]
oldsib4	-0.217*	-0.0702	-0.0141	0.00509
	[0.13]	[0.12]	[0.15]	[0.13]
Other family background variables				
nmoves	-0.0356***	-0.0360***	-0.0356***	-0.0290***
	[0.0085]	[0.0086]	[0.0085]	[0.0096]
faminc	7.01E-07***	5.48E-07	-1.12E-07	-7.6E-08
	[0.00000027]	[0.00000038]	[0.00000028]	[0.00000033]
famedsup	0.00599	0.00896	0.0271	0.0333
	[0.019]	[0.018]	[0.022]	[0.021]
twop	-0.202***	-0.272***	-0.165**	-0.217***
	[0.063]	[0.060]	[0.066]	[0.067]
singlep	-0.055	-0.00357	-0.054	0.0132
	[0.056]	[0.057]	[0.066]	[0.062]

Table 10 continued

	Dependent Variable			
Variable Name	**hedld3 (4 levels)**	**attain3 (11 levels)**	**hedld4 (4 levels)**	**attain4 (11 levels)**
Parental education variables				
nohs	-0.299***	-0.211***	-0.247**	-0.199**
	[0.082]	[0.081]	[0.099]	[0.089]
somepse	0.200**	0.0717	0.0852	0.0507
	[0.079]	[0.086]	[0.089]	[0.100]
privpse	0.0266	0.127	0.102	0.172*
	[0.11]	[0.091]	[0.12]	[0.099]
diploma	0.127***	0.0990*	0.108*	0.154***
	[0.047]	[0.052]	[0.055]	[0.050]
ucert	0.169**	0.262***	0.173*	0.409***
	[0.078]	[0.067]	[0.088]	[0.081]
bach	0.248***	0.371***	0.182***	0.473***
	[0.056]	[0.058]	[0.068]	[0.071]
prof	0.258**	0.612***	-0.0309	0.602***
	[0.12]	[0.13]	[0.092]	[0.12]
grad	0.373***	0.580***	0.179**	0.744***
	[0.078]	[0.076]	[0.077]	[0.071]
Parent's educational aspirations for child				
etrade	0.420***	[0.12]	0.635***	0.445***
	[0.12]	0.539***	[0.14]	[0.14]
ecoll	0.657***	[0.11]	1.027***	0.831***
	[0.11]	0.901***	[0.14]	[0.13]
euni	0.820***	[0.11]	1.146***	1.234***
	[0.11]	0.652***	[0.12]	[0.12]
epse	0.756***	[0.12]	1.217***	0.976***
	[0.12]	[0.12]	[0.14]	[0.14]
Trouble-makers among peers				
sometrbl	-0.0515	-0.113***	-0.00728	-0.0988***
	[0.038]	[0.037]	[0.035]	[0.035]
mosttrbl	-0.256***	-0.239***	-0.0914	-0.191**
	[0.075]	[0.075]	[0.085]	[0.089]
alltrbl	-0.250**	0.0848	-0.163	0.0669
	[0.12]	[0.14]	[0.14]	[0.13]

Table 10 continued

	Dependent Variable			
Variable Name	hedld3 (4 levels)	attain3 (11 levels)	hedld4 (4 levels)	attain4 (11 levels)
Province of schooling				
nfld	0.265**	0.285***	0.16	0.360***
	[0.11]	[0.093]	[0.10]	[0.090]
pei	0.17	0.336***	-0.0091	0.406***
	[0.11]	[0.12]	[0.10]	[0.11]
ns	0.126	0.409***	0.00146	0.439***
	[0.087]	[0.083]	[0.091]	[0.084]
nb	0.226**	0.294***	0.0435	0.334***
	[0.10]	[0.095]	[0.092]	[0.10]
que	0.503***	-0.0919	0.508***	-0.0376
	[0.092]	[0.077]	[0.091]	[0.069]
man	-0.146*	0.0485	-0.226**	0.0427
	[0.088]	[0.099]	[0.088]	[0.092]
sask	0.0314	0.0939	-0.0166	0.153*
	[0.095]	[0.090]	[0.090]	[0.091]
alta	-0.240***	-0.313***	-0.105	-0.118
	[0.084]	[0.091]	[0.081]	[0.073]
bc	-0.0485	-0.0341	-0.0299	-0.0118
	[0.072]	[0.078]	[0.076]	[0.082]
Size of community				
rural	0.127	-0.062	0.154*	-0.124
	[0.092]	[0.077]	[0.088]	[0.083]
stown	0.0059	-0.0118	-0.0175	-0.0509
	[0.064]	[0.063]	[0.065]	[0.066]
town	-0.138**	-0.123**	-0.101	-0.124**
	[0.062]	[0.060]	[0.062]	[0.060]
ccentre	0.00749	0.0017	0.0226	0.136*
	[0.12]	[0.090]	[0.086]	[0.080]
ncentre	-0.228**	-0.18	-0.0631	0.0196
	[0.090]	[0.11]	[0.073]	[0.081]
Characteristics of school-mates' families				
avefaminc	0.000000254	0.00000356**	0.00000107	0.00000325*
	[0.0000020]	[0.0000017]	[0.0000020]	[0.0000017]
puni	0.0766	-0.0702	0.0699	0.116
	[0.20]	[0.19]	[0.19]	[0.20]
ppemp	-0.251	-0.33	0.0421	0.351
	[0.48]	[0.41]	[0.48]	[0.41]

Table 10 continued

	Dependent Variable			
Variable Name	**hedld3 (4 levels)**	**attain3 (11 levels)**	**hedld4 (4 levels)**	**attain4 (11 levels)**
School characteristics				
schlsize	-0.000023	-2.6E-05	-0.0000016	2.82E-06
	[0.000049]	[0.000044]	[0.000052]	[0.000050]
coed	0.0228	-0.0816	-0.139	-0.0902
	[0.18]	[0.12]	[0.11]	[0.13]
private	0.459***	0.0232	0.475***	0.0746
	[0.11]	[0.083]	[0.11]	[0.087]
relschl	0.0492	0.0678	0.0497	0.158***
	[0.058]	[0.056]	[0.059]	[0.056]
stratio	-0.00667	-0.0076	-0.00605	-0.0181*
	[0.010]	[0.0094]	[0.0091]	[0.0096]
propqual	-0.401***	-0.182	-0.358***	-0.192
	[0.12]	[0.13]	[0.13]	[0.13]
scmatbui	-0.0475	-0.054	-0.0507	-0.0537
	[0.036]	[0.035]	[0.034]	[0.034]
scmatedu	0.0443*	0.0263	0.0578**	0.0481*
	[0.026]	[0.027]	[0.024]	[0.028]
skills	0.0913*	0.021	0.0565	0.0197
	[0.050]	[0.042]	[0.048]	[0.042]
ratcomp	0.0645	0.0929	-0.159	-0.101
	[0.24]	[0.23]	[0.30]	[0.21]
bullying	-0.0888	-0.121	-0.0632	0.0045
	[0.10]	[0.083]	[0.094]	[0.075]
drugs	-0.0805	-0.0641	-0.0394	-0.025
	[0.060]	[0.061]	[0.065]	[0.063]
poorhome	-0.0206	-0.00047	-0.0229	-0.0531
	[0.044]	[0.046]	[0.043]	[0.048]
tcmorale	-0.00999	0.00146	0.0289	0.0142
	[0.024]	[0.020]	[0.022]	[0.022]
schauton	-0.0134	-0.0145	-0.000594	0.0373
	[0.029]	[0.030]	[0.028]	[0.030]
tchparti	0.00251	-0.00013	0.0193	0.000985
	[0.019]	[0.020]	[0.019]	[0.021]

Table 10 continued

	Dependent Variable			
Variable Name	**hedld3 (4 levels)**	**attain3 (11 levels)**	**hedld4 (4 levels)**	**attain4 (11 levels)**
Threshold values				
mu1	-0.0224	1.863***	-0.763*	2.010***
	[0.50]	[0.42]	[0.45]	[0.46]
mu2	1.128**	1.935***	0.294	2.136***
	[0.50]	[0.42]	[0.45]	[0.46]
mu3	3.383***	2.888***	1.834***	3.163***
	[0.50]	[0.42]	[0.45]	[0.47]
mu4		2.995***		3.237***
		[0.42]		[0.47]
mu5		3.006***		3.253***
		[0.42]		[0.47]
mu6		3.039***		3.314***
		[0.42]		[0.47]
mu7		5.113***		5.537***
		[0.42]		[0.47]
mu8		5.152***		5.614***
		[0.42]		[0.47]
mu9		5.349***		5.827***
		[0.42]		[0.47]
mu10		5.541***		6.072***
		[0.42]		[0.47]
No. of obs.	8906	8914	7399	7419
Log pseudo L	-8053.045		-7731.731	
Wald χ^2 [p-value]	1934.25 [0.000]	2257.79 [0.000]	1180.10 [0.000]	2559.01 [0.000]
pseudo R-squared	0.1749	0.1596	0.1211	0.1770

Notes: The dependent variables in these equations indicate the highest level of education taken. The values in the table are maximum likelihood coefficient estimates, with robust standard errors corrected for clustering at the school level in square brackets. Estimation was carried out using sampling weights from the YITS-A parent questionnaire. One asterisk denotes statistical significance at the 10% level, two asterisks denote statistical significance at the 5% level, and three asterisks denote statistical significance at the 1% level.

Table 11. Marginal effects of school resources on the probability of post-secondary education

	Dependent Variable			
Explanatory Variable	**HEDLD3**	**HEDLD4**	**ATTAIN3**	**ATTAIN4**
Discrete change in probability				
private	0.18	0.19	*0.01*	*0.03*
relschl	*0.01*	*-0.05*	*0.03*	0.06
skills	0.03	*0.02*	*0.01*	*-0.01*
Derivative				
propqual	-0.15	-0.14	*-0.07*	*-0.07*
scmatedu	0.02	0.02	*0.01*	0.02
stratio	*-0.02*	*-0.02*	*-0.03*	-0.01

Notes: Marginal effects are derived from parameter estimates in Table 10. For dummy variables, the marginal effect measures the change in probability when the variable changes from 0 to 1. For continuous variables, the marginal effect is the derivative of the probability of a particular outcome. In all cases, marginal effects are summed across the outcomes that involve post-secondary education. Italics indicate that the marginal effect is related to a variable whose coefficient is not statistically significant at the 10% or lower level of significance.

Endnotes