AUSTRALIAN JOURNAL OF LABOUR ECONOMICS
Volume 14 • Number 3 • 2011 • pp 307 - 332

How Important are Omitted Variables, Censored Scores and Self-selection in Analysing High-school Academic Achievement?

Guyonne Kalb, University of Melbourne
Sholeh Maani, University of Auckland

Abstract

Using a rich longitudinal data set from birth, we explore three estimation issues related to academic performance analysis. Our paper primarily examines the effect of omitting childhood and teenage characteristics (childhood ability, parental resources at different times and peer effects), which are traditionally unavailable in data sets. Additionally, we explore the potential endogeneity of pre-exam school-leaving choices (self-selection) to academic performance; and we demonstrate the effect of accounting for censored academic performance measures. We find that omitting background characteristics results in overestimation of coefficients on other characteristics (the effect of current income is overestimated by 0.21 standard deviations of the average academic performance and the effect of ethnicity by 1.38 standard deviations). This then affects the policy implications drawn: for the group who did not take the exam, the predicted performance goes from a fail to a C (or pass). We also find that accounting for censored academic performance measures affects the estimation results, but allowing for selection correction does not.

JEL Classification: I21, J24, J13, J18

1. Introduction

A growing body of economic research is focusing on the academic performance of children and adolescents, as an important economic outcome of investments in education by families and communities. A problem identified across a wide range of advanced countries is that the proportion of young persons who leave high school without qualifications is alarmingly high. For example, on average across all

Address for correspondence: Associate Professor Guyonne Kalb, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne, Melbourne VIC 3010, Australia. Email: g.kalb@unimelb.edu.au
Acknowledgments: We are grateful to two anonymous referees, participants at the EALE/SOLE 2010 meeting and the Australasian Labour Econometrics Workshop meeting for comments. We are particularly indebted to Bruce Chapman for comments, and to David Fergusson and John Horwood for making the data set available and for information on the data.

OECD countries, 13 per cent of young people leave high school without secondary qualifications, and this percentage is higher than 20 per cent in some countries (OECD, 2008: p. 66).

In this paper we use a longitudinal data set from birth to provide evidence on the importance of including childhood family income and cognitive development, and teenage peer behavioural characteristics that are usually not available. An unresolved question is the relative importance of early childhood versus later income and background characteristics, as highlighted in the review by Haveman and Wolfe (1995) and in several studies since then. We observe a wide range of data characteristics for a cohort of students in New Zealand for whom we also observe Year 10 National Examination results if they took the exam. Importantly, our data set (Christchurch Health and Development Study (CHDS)) allows us to incorporate an extensive range of characteristics for a complete birth cohort, including students who do not have observed examination scores. We use this feature of our data to predict academic performance for all students (including those who did not take the exam) based on the students who took the exam while allowing for selection into taking the exam. We examine the effect of early childhood and later family income. We are also able to control for the effect of childhood cognitive development and teenage peers' behavioural effects in our predictions of expected grades for at-risk students. We show that our data is comparable to those in current studies when we use background characteristics that are readily available in the literature.

In New Zealand in general and in our study, students who are at school are expected to take School Certificate Exams. These are nationally administered exams, based on the same set of questions and grading for all participants, at the end of Year 10, usually at age 15. This is a great advantage as the use of this measure of academic performance eliminates recognised problems with inconsistency in comparing grades across schools, especially across lower and higher income school districts.[1] It thus provides nationally comparable academic performance results while in secondary school.

We contribute to the existing literature by testing the sensitivity of our academic performance results to the inclusion of a range of additional variables which are traditionally not available in many other data sets (childhood ability, parental resources at different times and peer effects). We find that the inclusion of these variables changes the coefficients on other variables, and alters the policy conclusions drawn from the results.

This is relevant for example when considering the legal school-leaving age. Such a policy is explored in many OECD countries as a means of improving the educational outcomes across population groups. Australia, the UK and New Zealand have, for example, recently considered raising the legal school leaving age. The minimum school-leaving age has been increased to 18 in, for example, Belgium and Netherlands and in some US states. The effectiveness of such a change relies heavily on the expected academic performance of students who are currently often not included in the relevant analyses due to early school leaving beyond the compulsory schooling age.

[1] After the year 2002, the School Certificate examination grades (also called the National Certificate of Educational Achievement (NCEA) examination) include a significant internal assessment component provided at the school level. Therefore, the current data provide a special opportunity to use nationally comparable School Certificate results before these changes were introduced.

In the context of academic performance analysis, and for students at risk of leaving high school early and without qualifications in particular, modelling issues of selection and censoring are pertinent. As a second contribution, this paper addresses both issues. Almost all academic performance data potentially have a self-selection aspect with regard to the personal choice of taking the examination. A question of interest is whether this selection feature of academic performance data, due to the choice of not taking the examination or dropping out of school beyond the compulsory schooling age, is expected to affect the estimation results for academic performance.

Few studies to date have addressed the selection question, and particularly in New Zealand study of this issue has been scarce. Card and Rothstein (2007) examine the effect of family background, and school and peer characteristics on the SAT score racial gap. They note the potentially problematic selection nature of SAT exam participation and mention unobserved pupil ability as the main problem in implementing a selection correction caused by lack of information for the students who do not take the SAT exam. Clark, Rothstein and Schanzenbach (2009) examine the effects of sample selection in school-level average SAT results in the US. Hansen, Heckman and Mullen (2004) examine the effects of sample selection in average Armed Services Vocational Aptitude Battery results in the National Longitudinal Survey of Youth (NLSY) in the US. The results from these papers do not unambiguously answer the question whether selection affects results. It remains unresolved across types of academic performance measures and institutional settings.

Furthermore, most academic performance data based on grade point averages (GPAs) are potentially affected by censoring at the lowest and the highest grade values of single subjects making up the GPAs. Cumulative and semester high-school and university GPAs are widely used as useful and comprehensive measures of academic performance, and their use is increasing across a number of recent studies in economics and related fields (e.g. Trockel, Barnes and Egget, 2000; Paschall and Freisthler, 2003; McKenzie and Schweitzer, 2001; Parker *et al.*, 2004; Stinebrickner and Stinebrickner, 2008). Therefore, examining the effects of allowing for censoring in GPA data is of interest. Only Hansen, Heckman and Mullen (2004) take censoring explicitly into account. In addition, censored data is potentially expected in major examinations that are designed to determine pass or fail outcomes. For example, major high-school examinations which determine suitability for continuation into academic or alternatively vocational high-school streams are likely to result in censoring at the lowest level of exam grades.[2] In the case of the New Zealand School Certificate exams used in our study, left-censoring, in particular, is evident.

In our analysis, we use a general specification that allows for censoring and selection in academic performance data. We jointly estimate the equation of academic performance, allowing for the censoring, with the equation for participation in the exam. Based on the school entry and school leaving rules in New Zealand, birth month is used as the exclusion restriction in our joint estimations.

The plan of the paper is as follows. A brief presentation of the analytical framework is provided in section 2. A discussion of the data set and the characteristics

[2] In fact, even in the case of examinations that are standardised at the national level (e.g. the SAT and average ACT scores) censoring of scores at the state or school levels is possible , due to higher or lower state-level or school-level achievement compared to the national-level score distribution.

of the sample follow in section 3. The estimated models and their results are presented in section 4. First, we explore the sensitivity of academic performance analysis to the exclusion of childhood and later background characteristics (childhood ability, parental resources and peer effects). Second, we explore the endogeneity of pre-exam school-leaving choices (self-selection) on potential academic performance. That is, we examine whether academic performance of those taking the exam is representative of the complete student population, including those who left school early. Third, we investigate the importance of allowing for censoring in individual subject scores. Conclusions are presented in section 5.

2. Model Specification and Estimation

We estimate models of academic performance (High-school examinations GPA) using Tobit analysis, jointly with exam participation choices (Probit), to account for censoring and potential selection effects (and comparable single-equation Tobit and OLS results on academic performance). We estimate academic performance for both students who took the exam and also importantly for out of sample at-risk students in our sample who had left school prior to examinations and identify potential contributing factors for at-risk students. We finally examine the sensitivity of our results to the inclusion of early childhood background and teenage peer characteristics available in our data set.

In the first subsection, we discuss the analytical framework that is used as the basis for the modelling of academic performance. The selection issue and the selection of an instrumental variable are discussed in the second and third subsection, followed by subsections on data censoring and the model specification.

Analytical Framework

Academic performance of children and adolescents has recently received significant interest in the literature in models that link parental and environmental resources to children's academic performance.[3] The recent literature on educational attainment has emphasised the significance of parental investments in human capital since childhood. An important implication of this literature is the recognition of the importance of academic performance and human capital investments throughout childhood on later schooling choices (see for example, Feinstein and Symons, 1999; Ermisch and Francesconi, 2001; Carneiro and Heckman, 2002; Haveman and Wolfe, 1995; Duncan *et al*., 1998; Gregg and Machin, 1998; Blau, 1999; Maani and Kalb, 2007; Rangvid, 2010; Sandy and Duncan, 2010).

Models of children's academic performance are usually based on a production function, where the parents are the producers. However, the academic performance of teenagers at the end of Year 10 and their school-leaving choices prior to the Year 10 exams are likely to be joint decisions, influenced, amongst other things, by the adolescent's academic ability and human capital investments by parents throughout childhood.

[3] Blau (1999), Borjas (1995), Case and Katz (1991), Duncan *et al*. (1998) and Montgomery (1991) for the US; Feinstein and Symons (1999) and Ermisch and Francesconi (2001) for the UK; and Miller and Volker (1989), and Prior and Beggs (1989) for Australia are examples of studies on the link between parental resources and educational attainment or labour market outcomes for their children.

We model academic performance, as measured by the average score (GPA) in a set of examinations as:

$$A_i = f(S_i, X_i, v_i) \qquad (1)$$

where A_i is the academic performance grade point average score of student i in National Examinations (representing ability and effort); S_i represents personal talents and abilities; and X_i is a vector of personal and parental resources, and environment; and v_ii represents the effect of unobserved factors, such as motivation.

Potential Self-selection

It is a common feature of academic performance analysis that results can only be based on those students who take the exam. Clark, Rothstein and Schanzenbach (2009) find that use of observed school mean SAT scores instead of latent scores understates the within-school variation in achievement. Therefore, when estimating an academic performance (average grade) equation allowance needs to be made for the selectivity of students who take the exam and whose average grade is therefore observed. This is relevant in the case of the School Certificate Exam in New Zealand, but also for measures in other countries, such as the SAT (Scholastic Aptitude Tests) or ACT (American College Testing) scores for college entrance in the US or the secondary school National Qualifications Framework's GCSE (The General Certificate of Secondary Education) in the UK.

In our case, we only observe the average grade for the students who continued with schooling beyond the compulsory level and who took the exam. The option of not taking the exam is chosen by all students who left school in an earlier period, but in addition there are some students who stayed at school but for other (unobserved) reasons did not take the exam.

The decision to participate in education beyond the compulsory age is intrinsically related to a number of factors as part of an inter-temporal decision (Becker, 1993). For example, investment in higher education is expected to result in higher returns for those with greater ability and a taste for lifetime labour force participation. In addition, keeping ability constant, a greater potential to finance education will lead to greater participation (Willis and Rosen, 1979; and Rice, 1987). Willis and Rosen's analysis utilised structural models and emphasised self-selection, and Rice's application utilised reduced-form models of participation and emphasised the effect of current financial constraints on school-leaving choices of males and females in secondary school. Neither study had observable variables on academic ability such as IQ or academic test scores. Micklewright's (1989) study used childhood math and reading scores, and parental current income in examining schooling choices at age sixteen, in the UK.

In this framework, selection into educational choices beyond the compulsory level for the ith student, such as participation in post-compulsory education or participation in examinations that lead to higher education, are influenced by V_{ij}, the utility of net expected present value of life-time earnings at each level of educational attainment j (E_j):

$$V_{ij} = V(E_j(S_i), X_i, u_i), j = 0, 1. \quad (2)$$

where S_i represents individual talents and abilities; and X_i represents observable personal and environmental characteristics, which determine the individual's tastes, expectations and financial constraints, and u_i represents unobservable heterogeneity. Our selection model of participation in the exam includes the effect of personal characteristics such as cognitive ability, peer and school effects, and parental income which may affect the choice of dropping out of school before the national exam and/or taking the exam.

Empirical estimation of the probability of enrolment at post-compulsory education (Pr PCE) is based on:

$$\text{Pr PCE observed} = \Pr[(V_{i1} - V_{i0} = g(S_i, X_i, u_i) > 0] \quad (3)$$

where vectors of observables S_i and X_i result in participation in education when the expected net benefit from education beyond the compulsory level ($V_{i1} - V_{i0}$) is positive. Equation 3 is estimated using a Probit specification.[4] In this paper the dependent variable is binary as to whether or not the respondent has results for at least one exam subject, as opposed to leaving school or not taking any subjects.

We use a data set, which includes a remarkably large number of relevant variables, thereby reducing the importance of the unobserved components. For example, we are able to control for both early childhood IQ and family income in early childhood and in teenage years, and teenage peers' school performance and characteristics.

Notwithstanding the richness of the data, we estimate the equations for academic performance and participation in the exam jointly to allow us to formally examine and account for any unobserved factors influencing both exam taking and academic performance. This is important when extrapolating results on exam takers to students at risk of leaving high school early and without qualifications.

Selection of the Instrumental Variable: Birth Month

The combination of starting age at school and compulsory schooling age laws provide an exclusion restriction in our model. Children in the CHDS are born within a five-month interval starting in April 1977 and ending in August 1977. All children are eligible to start school on their fifth birthday.

The academic year in New Zealand corresponds with the calendar year. That is, the school year starts in late January and ends in mid December. As a result, the first and the second semesters correspond with the first half and the second half of the calendar year respectively.

Children whose birth dates fall in semester 1 start school in the first semester

[4] Assuming that the net benefits conditional on S_i and X_i and their underlying characteristics are normally distributed and that G is a linear function of S_i and X_i, the expected net benefit would also follow the standard normal cumulative density function ($V_{i1} - V_{i0} \sim N(S_i'\beta + X_i'\gamma, \sigma^2)$, with β, γ and σ^2 constant across the population). Under these assumptions, Equation 3 can be estimated as a Probit model.

of the year they turn five and take the kindergarten year for the remainder of the year. They progress to grade 1, the next year. However, those whose birth dates fall in semester 2 start school in the second semester of the year they turn five, but since they have had less than one semester of kindergarten, they generally have to complete a full kindergarten year in the following year and start grade 1 a year later than those children born in the first semester. Therefore, children born in semester 1 are expected to be only 14 years old at the beginning of the academic year when they reach Year 10. As the minimum school-leaving age was 15, this group is legally bound to enrol at school in Year 10. The second group, born in semester 2, turns 15 in the latter part of their Year 9 and can legally leave school without a requirement to enrol in Year 10 at all. Those in the CHDS who are a few months older due to being born in the first semester have to enrol in Year 10 and therefore are more likely to be at school at the time of examination. Birth month is expected to affect school leaving before Year 10 (and thus exam taking), but not academic performance. This is verified empirically in our sample, and it is used to identify the models in this paper. That is, the equation for exam-taking contains birth month as an explanatory variable but this variable is not in the academic performance equation.

In our sample, 37.6 per cent of the births fell in the last part of the first semester, and the rest of the births fell in the early part of the second semester. Individuals from the two birth month groups have similar personal and background characteristics in our sample. Table A1 in the appendix provides a summary of mean characteristics by semester of birth in our sample. These mean characteristics show that the two groups were not statistically different in childhood IQ, family income, Year 10 exam grades, and school characteristics, as confirmed by t-tests across the two sub-samples. However, a higher percentage of those who were born in the second semester, and could legally leave school before entering Year 10, had left school compared to those who were born in the first semester. This characteristic of our data provides useful counterfactual observations of students who are comparable otherwise, except for the ability to leave school legally before the start of Year 10.

Angrist and Krueger (1991, 1992) were among the first to use the effect of quarter of birth and school-leaving laws on schooling choices in the US. The setting of their modelling approach is, however, different from ours in that they use quarter of birth as an instrument for years-of-education effects on earnings. The use of quarter of birth in earnings models has been questioned, if quarter of birth exerts a direct effect on academic performance. In addition, the question of quarter of birth being a weak instrument has been raised, in relation to the quarter of birth instrument in these models (e.g. Bound, Jaeger and Baker, 1995).

Hansen, Heckman and Mullen (2004) use birth month and laws of school-leaving age as an exclusion restriction for joint estimation of schooling and academic performance results for the US to adjust for selection. Our approach is generally comparable to Hansen, Heckman and Mullen's approach. The approach uses regression discontinuities due to schooling laws. The birth month instrument is required to be correlated with participation in education, but not to be directly related to academic performance. A few recent studies have specifically examined whether birth month and school entry age have direct long-term effects on academic performance. Different results arise from these studies. Elder and Lubotsky (2009) find that the effect of starting kindergarten earlier on achievement test scores is only significant for the first few months of kindergarten. Fredriksson and Ockert (2005) and Bedard and Dhuey (2006) find in contrast that older students perform better, but they also find that this effect decreases significantly over time although it does not completely disappear. Cascio and Schanzenbach (2007), and Black, Devereux and Salvanes (2008) find that there is no long-term relationship between birth month and school entry age, and academic performance in test scores.

While the international literature on the potential effect of birth month on schooling outcomes is varied, there appears to be consensus that there is little effect in the long run. This growing literature further indicates that results are likely to vary due to a host of institutional school-entry and school-leaving age laws. For example, in many countries such as in the US and New Zealand, students can leave school at age fifteen or sixteen. In other countries students are required to complete a specific school level, such as Year 10, regardless of age. These may have different implications in relation to study design. In addition, international practices vary in relation to early tracking of students into academic or vocational tracks based on academic performance, which may potentially lengthen any potential effects of school entry age on later academic performance. New Zealand does not have a student tracking system and students can legally leave school at a fixed age rather than a school level. Therefore, New Zealand regulations overall provide favourable conditions to address the questions of interest in this paper.

We confirm the validity of the birth-month variable as an exclusion restriction for joint estimations using identification through functional forms. This is consistent across single-equation estimates and joint model estimations. We find no evidence of a direct effect of the birth-month binary variable in the academic performance model. When included in our academic performance equation, the Birth Month_Semester 1 variable is insignificant. The log likelihood value for the jointly estimated model where the birth-month coefficient is restricted to zero in the academic performance equation is -655.68, versus a value of -655.52 for the full model including the variable. We cannot reject the restricted model. Whereas the model with both birth-month coefficients set to zero is rejected versus the model with an unrestricted birth-month coefficient in the sit exam equation at just over the 5 per cent-level of significance (a log likelihood value of -657.51 versus -655.68).

It is clear that the effect of birth-month on sitting the exam is relatively weak with a significance level of just below the 5 per cent-level. However, realistically the choices for a suitable instrument in these settings are limited, and the advantage of using the birth-month instrument to provide joint estimations of academic performance and participation in exams outweigh the disadvantage of the relatively weak instrument, which may at least be partly due to the relatively small sample size.

Censored Data

If academic performance, Ai, were a continuous measure, equation 1 could be estimated through Ordinary Least Square estimation. A Tobit-like approach is used in this case, because of the censored nature of exam grades below the Fail and above the grade A cut-offs. That is, the academic performance of the most capable and the least capable cannot be accurately measured through an exam targeted at the average student.

The dependent variable in the academic performance analysis is the average National School Certificate Examination grade, which is normally taken at age 15, on five subjects. For each subject, the score D for a fail is translated into a value of 0, a score C into 1, a score B into 2 and a score A into 3. Thus, the minimum score is 0 and the maximum score is 3. The average ranges in value between 3 for an A average to 0 for a D (fail) average. This specification is compatible with the official GPA (Grade Point Average) score assignment in New Zealand. While the GPA in this analysis is based on a 0-3 scale (a 4-point scale), it easily translates across any GPA scale practice, e.g. based on a 5-point or 10-point scale.

The dependent variable to be used in our analysis is a GPA and is constructed by averaging the numeric values of the score over all subjects taken in the certificate. In our data there is evidence of both left and right-censoring (10.62 per cent left-censored and 1.26 per cent right-censored). Left-censoring is significantly more pronounced given the nature of this type of examination which aims to determine suitability for continuation on an academic education path, as opposed to vocational or other paths. Individuals below or above a certain academic performance level in any of the subjects cannot be exactly ranked besides observing the minimum or the maximum level that they have achieved, depending on whether they had at least one score of 0 (a maximum level is observed and the observation is left censored) or of 3 (a minimum level is observed and the observation is right censored).

If only the average grade were available, we would estimate this model using a Tobit specification to allow for the censoring at 0 and 3. However, our data reports the grade for each subject. We would like to use the information on whether the student received a grade 0 or a grade 3 for at least one subject taken in the exam. For individuals who have at least one 0 score, we know that their underlying performance is less than the average score reported since 0 is the minimum value. For individuals who have at least one score of 3, we know that their underlying performance is more than the average score reported since 3 is the maximum value. The observations of individuals who have at least one score of 0 and also one score of 3 add no information since the latent academic performance could take any value. These observations are therefore dropped from the estimation. This reduces the sample from 601 to 578 respondents for the performance equation.

The model for latent academic performance $\overline{A}_i^*$ looks as follows:

$$\overline{A}_i{*} = \alpha_a + X_{ai}'\gamma_a + \varepsilon_a \qquad \text{with } \varepsilon_a \sim N(0, \sigma_a^2) \tag{4}$$

However, we can only compute the (GPA) average score $\overline{A}_i$ from $\overline{A}_{ij}$ (the score for subject *j* which lies in between 0 and 3), where $\overline{A}_{ij}$ is averaged over all subjects *j*. $\overline{A}_i$ may be censored at the lower or upper end:

$$\overline{A}_i = \overline{A}_i{*} \ \mathit{if} \ 0 < A_{ij}{*} < 3 \text{ for all subjects } j \text{ (indicator } d_i = 1) \tag{5}$$

$$\overline{A}_i \geq \overline{A}_i{*} \ \mathit{if} \ A_{ij}{*} \leq 0 \ (A_{ij} = 0) \text{ for at least one subject } j \text{ (indicator } d_i = 2) \tag{6}$$

$$\overline{A}_i \leq \overline{A}_i{*} \ \mathit{if} \ A_{ij}{*} \geq 3 \ (A_{ij} = 3) \text{ for at least one subject } j \text{ (indicator } d_i = 3) \tag{7}$$

This equation is estimated using maximum likelihood, taking the censoring into account:

$$\ln L = \sum_{\forall i,\ \text{if } d_i=1} \ln\left[\Pr\left(\varepsilon_a = \overline{A}_i - \alpha_a - X_{ai}{}'\gamma_a\right)\right] + \sum_{\forall i,\ \text{if } d_i=2} \ln\left[\Pr\left(\varepsilon_a \le \overline{A}_i - \alpha_a - X_{ai}{}'\gamma_a\right)\right] \quad (8)$$

$$+ \sum_{\forall i,\ \text{if } d_i=3} \ln\left[\Pr\left(\varepsilon_a \ge \overline{A}_i - \alpha_a - X_{ai}{}'\gamma_a\right)\right]$$

In addition to the Tobit-like specification in equation (8), we also examined alternative specifications. We compare results based on the Tobit-like specification to results obtained from simple OLS regressions on the average grade, ignoring the censoring completely, and to results from a Tobit regression on the average grade, ignoring the censoring occurring to individual subject grades. The specification in equation (8) utilises greater information on the latent score and unobservable factors such as ability, by using information on whether the grade was censored for any of the subjects included in the GPA. The OLS results are available in the appendix.[5]

Model Specification

Academic performance at the end of Year 10 is expected to be influenced by many personal, school and family resource variables, which influence school-leaving choices at age 15 as well. In addition, the same set of unobserved variables can potentially influence both academic performance and school leaving. With the use of the CHDS, we are able to control for a remarkably large number of important variables, which are often not observed in other data sets, reducing the unobserved component of the two equations. Our analysis allows us to examine the effects of family resources at different times, while controlling for important personal and peer-behavioural covariates such as childhood IQ and teenage peer effects.

We distinguish between parental income effects during early childhood and adolescent years on post-compulsory schooling choices and academic performance. Blau (1999) for example, provides evidence on the importance of permanent resources measures based on the NLSY data set. Rice (1987) and Micklewright (1989) show evidence of the impact of current income on schooling outcomes in the UK. Duncan *et al.* (1998) find evidence for the US based on the PSID data set that family economic conditions in early childhood are more pronounced determinants of completed schooling years than economic conditions later in life. Our two income measures are correlated, but the correlation is only 0.55 and thus each measure provides some independent information on the financial history of the household.

We also have a measure of the proportion of family income from welfare payments. The proportion of income from benefits was calculated based on data on all sources of parental welfare benefit income and other sources of income. The variable reflects the relative importance of welfare benefit income. The variable also reflects relative disadvantage with regard to the household's wealth and assets.[6]

---

[5] The Tobit results were fairly similar to the OLS results. They are not presented in this paper, but are available from the authors upon request.

[6] Rice (1987) used a 'current income' variable in addition to the 'benefit ratio' (the ratio of current benefit to current household income). In this study, the definition of the income and benefit variables is different from the Rice study, in that income is measured as the average family income decile between the ages of 11 to 14 and 1 to 5. The benefit ratio in this study is the only measure of current income (as proportion of household income when the child is aged 16). That is, the negative effect of the benefit ratio on school retention partly reflects the effect of economic disadvantage.

3. The Data

The Christchurch Health and Development longitudinal Study (CHDS) includes extensive economic and academic information on a cohort born in Christchurch, New Zealand, in 1977.[7] This cohort is followed throughout their childhood and adolescence, providing information on their transition from school to further education, training and work. The cohort includes all births in the city within a five-month period that started in April and ended in August. Christchurch, located in an English-speaking country, has a large population from the UK and other European backgrounds, and it is considered statistically representative of average New Zealand characteristics.[8] Among the advantages of this data set is the extensive amount of information on the cohort's academic and home environments, academic performance and ability, and socio-economic background.

The sample analysed in the study utilises information from survey years since birth to age 16 of the cohort. Our sample includes the respondents for whom data on all variables of interest were available. The original cohort of individuals in the survey consisted of 1265 individuals. The sample used in this study contains 713 observations for the analysis of school-leaving choices before the exams, 578 for the joint estimations of academic achievement at the end of Year 10 when using the Tobit-like specification explained in section 2.1.2.[9] The smaller sample based on students observed up to age 15 or 16 is partly due to minor attrition over time, and partly due to missing values on variables such as IQ, parental income and school factors. A detailed analysis of the data and comparisons with later Census data at both local and national levels show that the CHDS is fairly representative of families with children born around 1977.[10] In addition, our data includes all variables of interest for both the students who sat the exam, and those who did not due to school-leaving prior to exams.

The characteristics of this sample are summarised in table 1(a) below. Table 1(b) provides descriptions of all variables used in the analyses. Column 1 on the full sample shows academic performance and other characteristics that represent expected national averages, such as the average IQ of 102.8, and the average school certificate GPA of 1.19 (on a scale from 0 to 3) or a C.[11] Home ownership by parents was 88.6 per cent, and the average proportion of family income from benefits was 13.9 per cent. In the sample, 7.4 per cent had indigenous ethnicity (Maori) and 2.8 per cent were Pacific Islanders. In addition, 49.8 per cent of the mothers and 47.5 per cent of the fathers of

[7] For more information on the Christchurch health and Development Surveys longitudinal data set, the reader may refer to Fergusson, Horwood and Lloyd (1991), and Fergusson *et al*. (1989).
[8] Christchurch is the third largest city in New Zealand, an English speaking country. Christchurch has income and educational characteristics that resemble New Zealand national averages, but a higher proportion of the population (91.8 per cent) is from an English-speaking background, compared to the 80.1 per cent at the national level. The population of Christchurch and its surrounding areas is under half a million, and the other ethnic groups in the population include Maori, Pacific Island and other Asian and European ethnic groups.
[9] There are 601 observations in comparable OLS estimations of academic performance presented in table A2 in the appendix.
[10] A study of the CHDS for the New Zealand Treasury (Maloney, 1999) showed that attrition was related to some initial characteristics such as ethnicity and having a single parent. However, comparisons with later Census data at both local and national levels show that the CHDS is still fairly representative of families with children born around 1977.
[11] The average number of subjects taken was 5.09, with 52 per cent taking 5 subjects, 33 per cent taking 6 subjects and 6.9 per cent taking 4 subjects. On average, students had 0.62 A grades, 1.44 B grades, 1.72 C grades and 1.31 D grades.

the respondents had no school qualifications (less than the Year 10 School Certificate), and 20.6 per cent of mothers and 19.8 per cent of fathers had university qualifications.

Table 1(a) - Characteristics of the Sample

	Mean (Standard Deviation)	
Characteristics	*Full Sample*	*Took the National Exam*
Personal Characteristics		
Female (%)	50.5	52.6
	(50.0)	(50.0)
Indigenous Ethnicity (Maori) (%)	7.4	6.4
	(26.2)	(24.5)
Pacific Island Ethnicity (%)	2.8	2.6
	(16.5)	(15.9)
IQ (tested at 8 years of age)	102.8	104.8
	(14.89)	(13.66)
Education		
Average School Certificate Grade Point Average (GPA)	--	1.19
(taken in Year 10, with Fail=0, C=1, B=2, A=3)		(0.82)
Mother without Qualifications (<Year 10) (%)	49.8	45.8
	(50.0)	(49.9)
Mother with a Higher Education Qualification (%)	20.6	23.4
	(40.5)	(42.3)
Father without Qualifications (<Year 10) (%)	47.5	42.6
	(50.0)	(49.5)
Father with a Higher Education Qualification (%)	19.8	22.0
	(39.9)	(41.4)
Total Dropout rate from school at Age 16 (%)	9.8	--
	(28.4)	
Family and Social Environment		
Adolescent Average Income Decile: Ages 11-14		
(10 is most affluent Decile)	5.53	5.81
	(2.54)	(2.49)
Early Childhood Average Income Decile: Ages 1-5	5.82	6.09
	(2.40)	(2.32)
Own their Home (%)	88.6	91.7
	(31.8)	(27.6)
Number of Siblings	1.49	1.48
	(0.94)	(0.89)
Rural Location (%)	16.0	16.1
	(36.7)	(36.8)
Percentage of Family Income from Benefits (%)	13.9	11.0
	(32.8)	(29.2)
Local Unemployment Rate (%)	10.6	10.6
	(0.4)	(0.4)
Proportion of Respondent's class continuing to Year 11 (%)	83.7	85.8
	(16.2)	(11.6)
Average Class Size	28.8	28.9
	(4.20)	(4.13)
Association with Deviant Peers (10 is the highest association)	2.31	2.02
	(2.45)	(2.25)
Sample Size:	713	578

Table 1 (b) - Definition of Variables

Personal Characteristics	
Female	Binary=0 for a male; 1 for a female.
Ethnicity Indigenous (Maori)	Binary=1 if Maori.
Ethnicity (Pacific_Island)	Binary=1 if a Pacific Islander.
IQ8	The child's measured total Intelligence Quotient (IQ) score at 8 years of age (revised Wechsler Intelligence Scale for Children). This test is conducted by a trained Psychologist. (IQ score was not reported to the child, the parents or teachers).
Birth Month_ Semester 1	Binary=1 if born in April or May (school semester 1 in the southern hemisphere), =0 if born in June, July or August (school semester 2).
Education	
Average_Grade	The average value of all School Certificate (Year 10, age 15) examination marks over all subjects taken, with weightings of A =3, B =2, C =1 and 0 for a fail (D).
Mother_No_Qualifications	Binary=1 if child's mother does not have formal educational qualifications (Year 10 School Certificate or higher).
Mother_Higher_Qualifications	Binary=1 if a child's mother has a university or other higher education qualification.
Father_No_Qualifications	Binary=1 if a child's father does not have formal educational qualifications (Year 10 or higher).
Father_Higher_Qualifications	Binary=1 if child's father has a university or other higher education qualification.
Family and Social Environment	
Inc_Decile	Average income decile of the family when the child was between ages 11 and 14: 1 is consistently poor; 10 are consistently affluent.
Early_Inc_Decile	Average income decile of the family when the child was between ages 1 and 5: 1 is consistently poor; 10 are consistently affluent.
Own_Home	Binary=1 if parents own their own home and the child is living at home at 15 years of age.
Number_Siblings	Number of siblings in the home at 15 years.
Rural	Binary=1 if a child was not living in a main urban centre at 15 years of age.
Welfare_Benefit_Proportion	The proportion (between 0 and 1) of the family's income derived from social welfare benefits. This variable is expected to reflect relative disadvantage in terms of parental assets, relative income, and information or social networks.
Local_Unemployment	Unemployment rate by gender in the region in which each individual was living at 15 years of age. There were 8 regions. Corresponding levels of unemployment ranged between 5.9 and 12.1%. (Source: NZ Census, 1991, Regional Statistics).
Proportion of Class_Continue	Proportion of an individual's Year 10 (10th grade) class within the data set continuing onto Year 11 (11th grade). The relevant individual is excluded from the calculation.
Ave_Class Size	Average class size in secondary school
Peer_Deviant	Affiliation with deviant peers at age 15 based on self-reported use of tobacco, alcohol, illicit drugs, other illegal behaviour, etc. by friends in the previous year: scores range between 0 and 10, with 10 being the most deviant affiliations.

Column 2 of table 1(a) presents the mean characteristics for those who took the exam and reported at least one grade. In general, mean characteristic comparisons across columns 1 and 2 shows that those who participated in the exam had mean characteristics which were different from those of early school leavers. These differences include a higher average IQ at age 8, belonging to a higher family income decile, and going to a school with a higher proportion of the class continuing to Year 11.

4. Results

In this section, we discuss the results in four subsections. The model for academic performance is estimated taking into account self selection and censoring in the first subsection. We explore the expected academic performance of students who leave before taking the exam. We compare these expected results to the expected results of those taking the exam. The effect of omitting a number of often unobserved variables is reported next. The section concludes with a discussion of policy implications.

General Estimation Results

Table 2 presents the results for the joint model of academic performance and school leaving based on equations 4 and 3, allowing for the censoring of separate subject outcomes. The academic performance (Average_Grade) equation (the top section of table 2) shows that a number of personal and family factors since childhood (i.e. childhood IQ, parental income in the past, parent's education, and school and peer effects) are significant explanatory variables in determining academic performance.[12] In addition, girls perform better academically.

While the correlation between early childhood and the later income decile variable was 0.55, a large number of young adults in the sample had experienced changes in their family's income decile between early childhood and adolescent years. Some respondents experienced improvements in their household's relative income position, while others experienced deteriorations. Including both teenage and early childhood income decile variables, we find that they are both important in explaining academic performance. The effect of each higher recent income decile (averaged over the time when the respondent was aged between 11 and 14) is estimated to be equivalent to 0.160 of a full grade in the exam. In addition to this effect, early childhood income explains an additional effect of 0.159 of a grade per decile. Therefore, keeping other factors constant, together the predicted effect of the income decile variables is more than 0.3 of a grade difference for each income decile or close to a complete grade for 3 deciles difference (the difference between a C or a D average grade, for example). This does not take into account the censoring which would lower the predicted effect in terms of the GPA somewhat.

These results are consistent with the positive effect of income in many US studies (Haveman and Wolfe, 1995, and Duncan *et al*., 1998), and Gregg and Machin's (1998) findings on the effect of financial difficulties in early or late childhood. However, the finding by Duncan *et al*. (1998), that only early childhood parental income is significant in explaining the years of completed schooling, is not repeated

[12] In earlier versions of the model, we have included type of school (e.g. government, catholic, etc.), but this was found to be insignificant.

Table 2 - Academic Performance[a] (Tobit-like Specification 1)

	With Selection Equation			*No Selection Equation*		
Average_Grade (Tobit-like)	*Coefficient*	*z-value*	*P> |z|*	*Coefficiet*	*z-value*	*P>|z|*
Female	0.5577	3.14	0.002	0.5611	3.12	0.002
Ethnicity Indigenous (Maori)	0.2300	0.59	0.554	0.2275	0.61	0.542
Ethnicity (Pacific_Island)	-0.7785	-0.87	0.387	-0.7625	-1.09	0.275
Mother_No_Qualifications	-0.4492	-2.16	0.031	-0.4557	-2.13	0.033
Mother_Higher_Qualifications	0.2430	1.04	0.297	0.2414	1.01	0.315
Father_No_Qualifications	-0.0776	-0.39	0.698	-0.0894	-0.43	0.664
Father_Higher_Qualifications	0.3137	1.24	0.214	0.3086	1.21	0.226
Number_Siblings	0.0751	0.76	0.444	0.0751	0.73	0.467
Own_Home	-0.4881	-1.43	0.152	-0.4722	-1.33	0.184
Rural	0.2861	1.13	0.258	0.2878	1.08	0.278
Welfare Benefit_Proportion	0.6687	1.88	0.060	0.6581	1.70	0.089
Inc_Decile (ages 11-14)	0.1599	3.05	0.002	0.1594	3.18	0.001
Early_Inc_Decile (ages1-5)	0.1586	3.32	0.001	0.1596	3.25	0.0014
IQ8	0.0856	8.53	0.000	0.0869	10.08	0.000
Proportion of Class_Continue	2.1167	2.41	0.016	2.1905	2.54	0.011
Peer_Deviant	-0.1997	-4.58	0.000	-0.2047	-4.67	0.000
Ave_Class Size	0.0338	1.47	0.142	0.0344	1.46	0.143
Constant	-12.5656	-6.75	0.000	-12.8016	-8.04	0.000
Took the Exam (Probit)						
Female	0.1011	0.54	0.588			
Ethnicity Indigenous (Maori)	0.1157	0.44	0.661			
Ethnicity (Pacific_Island)	0.4904	1.15	0.249			
Mother_No_Qualifications	-0.2027	-1.11	0.265			
Father_No_Qualifications	-0.3055	-1.64	0.101			
Number_Siblings	0.0733	0.87	0.282			
Own_Home	0.3476	1.57	0.116			
Rural	0.0951	0.28	0.776			
Welfare Benefit_Proportion	-0.0381	-0.14	0.891			
Local_Unemployment	0.1757	0.67	0.505			
Inc_Decile (ages 11-14)	0.0123	0.23	0.816			
Early_Inc_Decile(ages1-5)	0.0329	0.82	0.412			
IQ8	0.0539	6.13	0.000			
Proportion of Class_Continue	2.0948	4.24	0.000			
Peer_Deviant	-0.1711	-5.62	0.000			
Ave_Class Size	0.0163	0.82	0.410			
Birth Month_Semester 1	0.3495	2.01	0.045			
Constant	-7.9176	-2.53	0.011			
Variance of the error term	1.6571	13.48	0.000	1.6556	--	--
Correlation (ρ)	-0.1302	-0.34	0.737			
No. of obs. = 713 (578 with obs. average grade)	Wald chi2(17)= 127.67			Number of observations = 578 LR chi2 (17) = 337.42		
Log pseudo likelihood = -655.6818	Prob > chi2 = 0.000			Log likelihood = -513.2106 Prob > chi2= 0.000		
Left-censored obs. = 292	Uncensored = 142			Right-censored = 144		

Note: a) Academic Performance: Average Grade in the National Examination in Year 10 (10th grade) Selection Equation: 1=Took the exam; 0=Left school before the exam or did not take exam.

here. Our results are consistent with results for the UK (Feinstein and Symons, 1999; and Ermisch and Francesconi, 2001) regarding the importance of resources throughout childhood in determining children's academic performance.

Childhood cognitive ability is also important. Each additional unit of childhood IQ score is associated with 0.086 of the Year 10 exam GPA. This is a large effect considering the range of IQ scores. The mean IQ score was 102.8 with a standard deviation of 14.8, a minimum of 70 and a maximum of 143. This effect highlights the importance of the respondent's childhood scholastic ability in predicting academic performance in later years. Comparing this effect to the combined effect of the two income decile variables, we see that one income decile in the Tobit-like estimates is equivalent to about 3.7 IQ units. (In the OLS, one income decile is equivalent to just over 3.2 IQ units). In addition, the mother's lack of school qualifications, and class and peer effects are significant in explaining academic performance. Peer effects measure affiliation with deviant peers at age 15 based on self-reported use of tobacco, alcohol, illicit drugs, other illegal behaviour, etc. by friends in the previous year: scores range between 0 and 10, with 10 being awarded to students with the most deviant peer affiliations. A move from the lowest to the highest score is equivalent to a drop of just over 23 IQ units or a drop in income in childhood and teenage years of just over 6 deciles.

The variable birth month, which is used as an instrument in the exam-taking equation, is significant and has the expected sign. Those who cannot leave school legally before the exam are more likely to stay at school until after the exam (and therefore sit the exam).[13] The other results of the exam-taking Probit indicate that, keeping other factors constant, children of parents without qualifications and children who associate with deviant peers are less likely to stay at school until after the exam. Children of parents who own their home, children who experienced higher childhood family income and children who had a higher IQ at age eight are more likely to stay at school. Once IQ and peer effects are included in the model, relatively few other variables are significant. For example, the variables for ethnic background and most indicators for parental education were insignificant.

The correlation coefficient ρ is small positive and insignificant, indicating that self-selection into taking the exams as a result of unobserved factors is not a major issue. Comparing the academic performance equation in the joint model with the equation for taking the exam to a single-equation model for academic performance (also in table 2), it is clear that the results in both models are very similar. This is as expected, given the small insignificant value for the correlation coefficient.

Using a comparable specification to earlier studies (a single equation for academic performance estimated by OLS), we obtain the results presented in table A2 in the appendix.[14] The much larger coefficients in column 4 of table 2 in our Tobit-like analysis reflect the fact that in the OLS specification, academic performance outcomes are restricted to lie in between 0 and 3, whereas in the Tobit-like specification, we allow

---

[13] There is a second variable in our model, which is included in the Probit selection equation only. The local unemployment rate is expected to affect school leaving directly due to job opportunities, but the unemployment rate is not expected to affect academic performance.

[14] We confirmed our conclusion regarding selectivity based on this alternative specification for the academic performance equation combined with the same selection equation as in table 2, using a two-step Heckman approach (Heckman, 1979), and find similar results regarding selectivity. It shows that the coefficient on the Heckman correction term in the academic performance equation is insignificant and that the other parameters change slightly only. These results are available from the authors.

a wider range of values as outcomes on the latent variable for academic performance. An important difference between the two approaches is that in the Tobit-like approach, the contribution to the log-likelihood of students with a GPA of X who had at least one score of 0 or 3 is different from the contribution to the log-likelihood of students with a GPA of X who had only scores of 1 and 2. Equation 8 shows that there is more uncertainty regarding the academic performance of the former type of students. This reflects the wider range of ability represented by students who score a 0 or 3. The larger (negative and positive) coefficients also indicate that characteristics that have a negative effect on academic performance are more likely to be associated with having at least one score of 0, and reversely, characteristics that have a positive effect on academic performance are more likely to be associated with having at least one score of 3. This reinforces the negative and positive effects estimated using OLS. Our Tobit-like estimates show that the effect of income for at-risk students may potentially be under-estimated by OLS. However, the direction and relative size of the effects is similar between the two specifications.

Prediction of Grades (GPA) for School Leavers

To further explore the implications of the estimated models with and without allowing for correlation in unobserved heterogeneity between exam taking (staying at school) and academic performance, we use the estimated parameters in table 2 to predict academic performance for everyone in the sample. The third row shows predictions based on OLS results from table A2.

Independent of the approach taken, we find that those who did not take the National Exam were expected to perform poorly and significantly below those who did take the National Exam (see table 3). This lower predicted score was mostly due to a difference in observable characteristics between the two groups.

Table 3 - Predicted Average Grade for Full Sample, Based on Specification 1 with and without Controlling for Sample Selection (as presented in tables 2 and A2)

Variable	*Mean*	*Std. Dev.*	*Minimum*	*Maximum*
A: Group 1: Took the exam				
Average_Grade allowing for censoring and selection[a]	1.015	1.089	0.000	3.000
Average_Grade allowing for censoring	1.010	1.091	0.000	3.000
OLS Average_Grade[b]	1.204	0.569	-0.147	2.876
B: Group 2: Dropped out before exam, or did not take the exam				
Average_Grade allowing for censoring and selection[a]	0.063	0.278	0.000	1.874
Average_Grade allowing for censoring	0.061	0.271	0.000	1.871
OLS Average_Grade[b]	0.265	0.498	-0.905	1.562

Notes: a) The rows that allow for censoring are based on the results in table 2, whereas the OLS rows are based on the results in table A2. The former are based on predictions which are censored at 0 and 3 to make the results more comparable to the OLS results. The latent predicted values for the former are 0.674 and 0.653 in group 1 and -2.379 and -2.462 in group 2. This indicates that although the values are quite different from the censored results, it remains clear that there is a difference in performance between the group that took the exam and the group that did not. b) The OLS predicted mean value corresponds to the sample mean value of the 601 observations on which the OLS estimations are based.

The GPA predictions for those who took the national exam (table 3, section A) show that these students had a predicted average grade of C (1.02 or 1.20) compared to a predicted grade of D (0.063 or 0.265) for those who did not take the exam (table 3, section B). Controlling for sample selection changes the results only slightly. The results from the different approaches for those who took the exam were quite similar. The largest change is for those who did not take the National Exam and dropped out of school prior to exams. The results all imply that those who had dropped out of school before taking the national exam were expected to perform poorly in the exam, and importantly this seems mostly explained by a difference in observed characteristics rather than unobservable characteristics.

Omitted Variables

As a sensitivity check, we leave out four important variables: IQ at age 8, family income decile in childhood, peer characteristics and the proportion of the class mates which continues beyond Year 10. Table 4 shows that once these important (and often unobserved) variables are left out, some of the coefficients change substantially. For example, having a Pacific Islander ethnicity now has a significant negative effect on performance. That is, the coefficient has increased significantly by 1.38 standard deviations of the average academic performance (from -0.7785 to -1.9069), indicating that ethnicity is now explaining some of the variation due to one or more of the omitted variables. We will get back to this issue in the section 4.4 on policy implications. In addition, the coefficient for recent family income increases significantly. Here, the coefficient has increased in size by 0.21 standard deviations of the average academic performance (from 0.1599 to 0.3291). Similar larger estimated effects are observed for the OLS results, when we exclude these four variables, in table A2 in the appendix.

Comparing these results to the analyses in the current academic performance literature, we find that they are comparable on the variables of interest that are available in the literature. We find that the OLS general specifications with our data (CHDS), which incorporate variables generally used in the literature (i.e. ethnicity, gender, locality, current household income, parental education, and school type) result in coefficients that are comparable to those in the current literature. For example, among a group of current studies considered (Sandy and Duncan, 2010 (NLSY 97); Rangvid, 2010 (PISA 2003), Stinebrickner and Stinebrickner, 2008), the effect of ethnicity on academic performance ranges between –0.25 to –0.80 SD (standard deviations of the academic performance measure). The equivalent OLS estimation with the CHDS data, based on a set of variables as specified above, results in an ethnicity effect within the range found in the current literature (–0.37 SD for Pacific Island ethnicity). This effect is comparable in size to the effect for urban Hispanic ethnicity (–0.41 SD) in Sandy and Duncan (2010), and for Blacks (–0.34 SD) in Stinebrickner and Stinebrickner (2008).

Likewise, the equivalent estimate for mother's higher education based on our data is within the range found in the literature, at +0.30 SD. The estimated effect is for example, +0.13 SD of academic performance for mother's some college education in Sandy and Duncan (2010), effects of +0.20 to + 0.23 SD of mother's higher occupation in Rangvid (2010), and 1.47 SD for mother's completion of a higher degree (sample of white males) in Hansen, Heckman and Mullen (2004)).

Table 4 - Academic Performance[a] (Tobit-like Specification 2: with Fewer Explanatory Variables)

	With Selection Equation			*No Selection Equation*		
Average_Grade (Tobit-like)	*Coefficient*	*z-value*	*P> |z|*	*Coefficiet*	*z-value*	*P>|z|*
Female	0.4091	1.83	0.067	0.4163	1.88	0.060
Ethnicity Indigenous (Maori)	0.0344	0.07	0.946	0.0306	0.07	0.948
Ethnicity (Pacific_Island)	-1.9069	-1.94	0.052	-1.8932	-2.05	0.041
Mother_No_Qualifications	-0.6378	-2.37	0.018	-0.6503	-2.44	0.015
Mother_Higher_Qualifications	0.7207	2.50	0.012	0.7197	2.38	0.017
Father_No_Qualifications	-0.3207	-1.21	0.226	-0.3415	-1.33	0.183
Father_Higher_Qualifications	0.8264	2.74	0.006	0.8239	2.61	0.009
Number_Siblings	0.0314	0.26	0.796	0.0303	0.24	0.810
Own_Home	-0.0663	-0.15	0.881	-0.0232	-0.05	0.958
Rural	0.4094	1.26	0.209	0.4095	1.25	0.210
Welfare Benefit_Proportion	0.6753	1.37	0.170	0.6515	1.39	0.164
Inc_Decile (ages 11-14)	0.3291	5.29	0.000	0.3320	5.50	0.213
Early_Inc_Decile(ages1-5)	--			--		
IQ8	--			--		
Proportion of Class_Continue	--			--		
Peer_Deviant	--			--		
Ave_Class Size	0.0065	0.22	0.823	0.0068	0.24	0.810
Constant	-1.7126	-1.56	0.119	-1.7926	-1.71	0.087
Took the Exam (Probit)						
Female	0.0988	0.61	0.542			
Ethnicity Indigenous (Maori)	0.0217	0.09	0.928			
Ethnicity (Pacific_Island)	0.2492	0.62	0.534			
Mother_No_Qualifications	-0.3124	-2.06	0.039			
Father_No_Qualifications	-0.4865	-3.23	0.001			
Number_Siblings	-0.0292	-0.42	0.671			
Own_Home	0.6518	3.60	0.000			
Rural	0.1997	0.66	0.510			
Welfare Benefit_Proportion	-0.2193	-1.01	0.314			
Local_Unemployment	0.2586	1.01	0.312			
Inc_Decile (ages 11-14)	0.0838	2.23	0.025			
Early_Inc_Decile(ages1-5)	--					
IQ8	--					
Proportion of Class_Continue	--					
Peer_Deviant	--					
Ave_Class Size	0.0101	0.57	0.570			
Birth Month_Semester 1	0.3043	2.09	0.036			
Constant	-2.3873	-0.85	0.098			
Variance of the error term	2.1802	13.02	0.000	2.1786	--	--
Correlation (ρ)	-0.0810	-0.43	0.669			
Number of observations = 713 (578 with observed academic performance)	Wald chi2(12) =89.50			Number of observations = 578		
Log likelihood = -812.7124	Prob > chi2 = 0.0000			Log likelihood = -603.7017		
Left-censored obs. = 292	Uncensored = 142			Right-censored = 144		

Note: a) Academic Performance: Average Grade in the National Examination in Year 10 (10th grade) Selection Equation: 1=Took the exam; 0=Left school before the exam or did not take exam.

Using the results from table 4 and table A2 to predict average scores in table 5, it is clear that leaving out these four important variables also has an effect on our ability to predict academic performance, particularly for those who did not take the exam.[15] The OLS results for the group who did not take the exam have clearly increased to a near pass.[16] When a number of important variables are not available and therefore cannot be included in the model, the model's predictions would erroneously indicate that early school leavers are expected to perform only slightly worse than students who are already taking the National Exam.[17] That is, they are expected to score a C instead of a D (Fail). As table 5 shows, the Tobit-like model performs better than the OLS, in predicting low GPAs for those students who are not taking the exam when a number of important variables are excluded.

Table 5 - Predicted Average Grade for Full Sample, Based on Specification 2 with and without Controlling for Sample Selection (as presented in tables 4 and A2)

Variable	*Mean*	*Std. Dev.*	*Minimum*	*Maximum*
A: Group 1: Took the exam				
Average_Grade allowing for censoring and selection[a]	0.854	1.013	0.000	3.000
Average_Grade allowing for censoring	0.843	1.014	0.000	3.000
OLS Average_Grade[b]	1.204	0.426	0.165	2.260
B: Group 2: Dropped out before exam, or did not take the exam				
Average_Grade allowing for censoring and selection[a]	0.197	0.489	0.000	2.225
Average_Grade allowing for censoring	0.188	0.481	0.000	2.211
OLS Average_Grade[b]	0.852	0.287	0.195	1.720

Notes: a) The rows that allow for censoring are based on the results in table 4, whereas the OLS rows are based on the results in table A2. The former are based on predictions which are censored at 0 and 3 to make the results more comparable to the OLS results. The latent predicted values for the former are 0.606 and 0.575 in group 1 and -0.503 and -0.568 in group 2. This indicates that similar to the censored results, the predicted performance of the group that took the exam has become much more similar to the predicted performance of the group that did not take the exam compared to table 3. b) The OLS predicted mean value corresponds to the sample mean value of the 601 observations on which the OLS estimations are based.

[15] The predicted OLS average grade for the group who took the exam has not changed across tables 3 and 5 since the predicted OLS average is equal to the observed average score by definition.
[16] Additional sensitivity analyses on the relative importance of these variables showed that the omission of each resulted in an increase in the predicted GPA. The largest effect was from omitting childhood IQ, which increased the predicted OLS GPA for the group of students who did not take the exams to 0.61 (0.26 based on the joint Tobit-like specification). When the two childhood variables (childhood IQ and early childhood income) were omitted the OLS GPA prediction was 0.62 (and 0.27 based on the joint Tobit-like specification), compared to a predicted GPA of 0.45 when the two teenage behavioural and peer variables were omitted (0.12 based on the joint Tobit-like specification). This highlighted the importance of early childhood characteristics in the analysis of students at risk of leaving school early.
[17] We checked and found that a two-stage selection adjustment did not remedy the over-prediction of the GPA for at-risk students based on the OLS model with omitted variables.

Policy Implications

Our results including all variables indicate that the students who had chosen to leave had a predicted grade of D, or fail, in the National Exam. Our results highlight that keeping students at school may be a useful first step, but it is not expected to be sufficient as a policy without other complementary interventions. That is, the students who do not participate in the National Exam are on average predicted to fail the exam based on their individual and environmental characteristics in childhood and as a teenager.

In addition, our results have policy implications in identifying family, school and peer characteristics that contribute to academic attainment. For example, our results lend support to the importance of early childhood family resources and early child development for the improvement of academic performance of population groups. It appears important to support children in families with limited resources from early childhood onwards to ensure good education outcomes from the start. In addition, the teenage peer behavioural variables are shown to be significant covariates of school-leaving choices and academic performance. This indicates that policies targeted at the school level rather than at the individual level may be more effective by changing outcomes for a peer group rather than focussing on one individual in the peer group. These behavioural aspects may be influenced through policy on learning and behavioural aspects of teenagers' school and peer environments. For example, Rothstein (2004) considers the SAT score in the US in the prediction of college performance, as opposed to high-school GPA as a predictor. Including information on the student's gender and ethnicity, and the (fraction) ethnicity of the high-school attended, he finds that the SAT score's role in educational prediction models may be quite sensitive to the inclusion of background variables. This is particularly true for high-school peer ethnicity characteristics as predictors. In that context, it is interesting to note that in our results once personal, socio-economic and environmental characteristics are controlled for, indigenous (Maori) and Pacific Island teenagers do not perform more poorly and they do not have a higher statistically significant probability of leaving school before the exam. This result based on our formal estimation is in agreement with Card and Rothstein's (2007) assertion that in general the absence of schoolmate characteristics would potentially lead to overestimation of negative ethnicity effects on academic achievement. A sensitivity analysis leaving out four important variables (as described in section 4.3) shows that this has substantial implications for the estimated coefficients and predicted academic performance.

Our results which show an increased ethnicity coefficient are consistent with Rothstein's (2004) results on ethnicity, indicating ethnicity serves as a proxy for factors such as school-peer achievement or peer behaviour, as opposed to ethnicity itself. In addition, we show the importance of childhood economic resources and cognitive development. These results show the relevance of identifying these underlying factors for effective policy design. For example, a policy that is targeted at ethnicity and peer effects is consistent with school desegregation, while a policy which also focuses on underlying causes would focus on the importance of childhood cognitive development and continued enhanced academic learning. Potential policies are childhood and later provision of assistance with homework for disadvantaged children and adolescents, or ensuring the availability of a quiet location to do homework. A growing current literature in education, for example, shows positive effects of pre-school education on childhood cognitive development, which benefits children from low-income and immigrant families in particular. The extent of the effects depends on the quality of the program and whether it is sustained beyond pre-school (e.g. see Gillian and Zigler, 2001; Boocock, 1995; and Schweinhart *et al*., 2005).

The results on predicted academic performance in section 4.3 show that when a number of important variables are omitted, the OLS model's predictions would erroneously indicate that policies aimed at keeping students at school are likely to be sufficient to improve their outcomes. The restricted model predicts that early school leavers would perform only slightly worse than students who are already taking the National Exam. Our analysis shows that the models allowing for censoring perform better than the OLS models when omitting important variables, regardless of controlling for selection.

5. Conclusion

In this paper, we have examined the effect of childhood and later family resources on the academic performance of high-school students, with a focus on expected outcomes for students at risk of early school leaving without qualifications. Our analysis shows that those who left school early were predicted to perform with a fail grade on average in the Year 10 National Exam. Academic performance can be predicted using the model estimated based on the students who are taking the exam. The results change considerably when some individual and environmental characteristics (such as childhood IQ, early childhood family resources, teenage association with deviant peers, and the proportion of the student's Year 10 who continue to Year 11) are excluded from the model. These variables are often not available for analysis. Using the restricted model leads to much less differentiation in predicted results for the two groups of students. We find that when these variables are excluded, recent family income and ethnicity serve as proxies for these omitted variables. That is, the effect of current income is overestimated by 0.21 standard deviations of the average academic performance and the effect of being from Pacific Islander descent is overestimated by 1.38 standard deviations.

In addition, we find that once a number of important variables are excluded, the model (OLS in particular) overestimates the potential performance of students who did not take the exam. That is, for this group, the predicted performance when all variables are included is a fail, whereas it turns into a C (or pass) when we exclude these variables. Our analysis shows that the model allowing for censoring performs better in this regard, regardless of controlling for selection.

In addition, the inclusion of individual, family and teenage peer information has the potential to influence the conclusions drawn from studies. For example, our results based on the full set of variables (table 2) show that in order to improve the educational outcomes of the student population, just keeping students at school by increasing the legal school-leaving age is not sufficient as a policy in isolation. This conclusion is based on the result that the academic performance of those who left prior to exams was predicted to be much poorer than the performance of those who took the exam. This poor result was driven by a number of observed characteristics (allowing for unobserved heterogeneity did not make a significant difference to the results). Some of these observable variables in our model reflect factors that have been in effect since early childhood. For example, both early childhood and later family resources are important covariates.

Finally, it was shown that if the (often unobserved) family, school and peer characteristics, (mentioned above) cannot be included, other variables such as school ethnicity or current income often serve as a proxy for these factors. Therefore, identifying the underlying factors, so they can be targeted (starting in early childhood and throughout teenage years), is important to enable the design of appropriate and effective policies.

Appendix

Table A1 - Childhood IQ, Family Income and Grade Characteristics by Birth-Month Sample

	Mean (Standard Deviation) Birth Month Falls in:	
Characteristics	*Semester 1*	*Semester 2*
Personal Characteristics		
IQ (tested at 8 years of age)	103.52	102.32
	(14.31)	(15.22)
Average_Grade (School Certificate Exam GPA in Year 10;	1.16	1.23
where Fail=0, C=1, B=2, A=3)[a]	(0.75)	(0.85)
Adolescent Average Income Decile: Ages 11-14 (10 is most	5.43	5.59
affluent Decile)	(2.45)	(2.60)
Early Childhood Average Income Decile: Ages 1-5	5.76	5.86
	(2.38)	(2.41)
Number of siblings	1.47	1.50
	(0.96)	(0.93)
n their Home	89 %	88 %
Mother had completed Higher Education	19 %	21 %
Father had completed Higher Education	20 %	20 %
Left school before the exam	7 %	10 %
Sample Size:	268	445

Note: a) The average score in the first column is based on 235 observations and the average score in the second column is based on 366 observations.

Table A2 - Academic Performance[a] (OLS Specifications)

	Specification 1			Specification 2		
Average_Grade	*Coefficient*	*t-value*	*P> |t|*	*Coefficiet*	*t-ratio*	*P>|t|*
Female	0.2040	4.14	0.000	0.1569	2.70	0.007
Ethnicity Indigenous (Maori)	0.0542	0.53	0.596	0.0415	0.34	0.731
Ethnicity (Pacific_Island)	-0.1238	-0.78	0.437	-0.3081	-1.64	0.102
Mother_No_Qualifications	-0.1279	-2.17	0.031	-0.1775	-2.54	0.011
Mother_Higher_Qualifications	0.1261	1.83	0.068	0.2481	3.05	0.002
Father_No_Qualifications	-0.0332	-0.58	0.564	-0.1108	-1.63	0.103
Father_Higher_Qualifications	0.1306	1.79	0.073	0.2831	3.37	0.001
Number_Siblings	0.0145	0.52	0.601	-0.0028	-0.09	0.931
Own_Home	-0.1076	-1.12	0.265	0.0161	0.14	0.888
Rural	0.0643	0.90	0.370	0.0766	0.91	0.365
Welfare Benefit_Proportion	0.2309	2.28	0.023	0.1770	1.48	0.140
Inc_Decile (ages 11-14)	0.0558	4.04	0.000	0.0957	6.40	0.000
Early_Inc_Decile(ages1-5)	0.0347	2.65	0.008	--		
IQ8	0.0270	14.14	0.000	--		
Proportion of Class_Continue	0.6053	2.67	0.008	--		
Peer_Deviant	-0.0558	-5.04	0.000	--		
Ave_Class Size	0.0064	1.01	0.312	0.0009	0.12	0.903
Constant	-2.8139	-7.89	0.000	0.5104	1.91	0.056
No. of Observations=601	$\bar{R}^2 = 0.475$			$\bar{R}^2 = 0.259$		

Note: a) Academic Performance: Average Grade in the National Examination in Year 10 (10th grade.

References

Angrist, J.D. and Krueger, A.B. (1991), 'Does Compulsory School Attendance Affect Schooling and Earnings?', *The Quarterly Journal of Economics*, 106, 979-1014.

Angrist, J.D. and Krueger, A.B. (1992), 'The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples', *Journal of the American Statistical Association*, 87, 328-336.

Becker, GS (1993), *Human Capital*, Third Edition, Chicago University Press.

Bedard, K. and Dhuey, E. (2006), 'The Persistence of Early Childhood Maturity: International Evidence of Long-run Age Effects', *Quarterly Journal of Economics*, 121, 1437-1472.

Black, S.E., Devereux, P.J. and Salvanes, K.G. (2008), 'Too Young to Leave the Nest? The Effect of School Starting Age', IZA Discussion Paper No. 3452.

Blau, D.M. (1999), 'The Effect of Income on Child Development', *Review of Economics and Statistics*, 81, 261-276.

Borjas, G.J. (1995), 'Ethnicity, Neighbourhoods, and Human-Capital Externalities', *American Economic Review,* 85, 365-390.

Boocock, E.F. (1995), 'Early Childhood Programs in Other Countries: Goals and Outcomes', *The Future Children*, 5, 25-50, Los Altos.

Bound, J., Jaeger, D.A. and Baker, R.M. (1995), 'Problems with Instrumental Variable Estimation when the Correlation between the Instrument and the Endogenous Explanatory Variable is Weak', *Journal of the American Statistical Association*, 90, No. 430, 443-450.

Card, D. and Rothstein, J. (2007), 'Racial Segregation and the Black-White Test Score Gap', *Journal of Public Economics*, 91, 2158-2184.

Carneiro, P. and Heckman, J.J. (2002), 'The Evidence on Credit Constraints in Post-secondary Schooling' *Economic Journal,* 112, 705-734.

Cascio, E. and Schanzenbach, D.W. (2007), 'First in the Class? Age and the Education Production Function', NBER Working Paper 13663.

Case, A.C. and Katz, L.F. (1991), 'The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths', NBER Working Paper 3705.

Clark, M., Rothstein, J. and Schanzenbach, D.W. (2009), 'Selection Bias in college Admission Test Scores', *Economics of Education Review*, 28, 295-307.

Duncan, G.J., Yeung, W.J., Brooks-Gunn, J. and Smith, J. (1998), 'How Much Does Child Poverty Affect the Life Chances of Children?', *American Sociological Review*, 63, 406-423.

Elder, T.E. and Lubotsky, D.H. (2009), 'Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background and Peers', *Journal of Human Resources*, 44, 641-683.

Ermisch, J. and Francesconi, M. (2001), 'Family Matters: Impact of Family Background on Educational Attainment', *Economic*a, 68, 137-156.

Feinstein, L. and Symons, J. (1999), 'Attainment in Secondary School', *Oxford Economic Papers*, 51, 300-321.

Fergusson, D.M., Horwood, L.J. and Lloyd, M. (1991), 'Confirmatory Factor Models of Attention Deficit and Conduct Disorder', *Journal of Child Psychology and Psychiatry*, 257-274.

Fergusson, D.M., Horwood, L.J., Shannon, F.T. and Lawton, J.M. (1989), 'The Christchurch Child Development Study: A Review of Epidemiological Findings', *Paediatric and Perinatal Epidemiology*, 3, 302-325.

Fredriksson, P. and Ockert, B. (2005), 'Is Early Learning Really More Productive? The Effect of School-starting Age on School and Labor Market Performance', IZA Discussion Paper No. 1659.

Gregg, P. and Machin, S. (1998), 'Child Development and Success or Failure in the Youth Labour Market', Centre for Economic Performance Discussion Paper: no 397, London School of Economics.

Gillian, W.S. and Zigler, E.F. (2001), 'A Critical Meta-analysis of all Evaluations of State-funded Preschool from 1977 to 1998. Implications for Policy, Service Delivery and Program Effectiveness', *Childhood Research Quarterly*, 15, 441-473.

Haveman, R. and Wolfe, B. (1995), 'The Determinants of Children's Attainment: A Review of Methods and Findings', *Journal of Economic Literature*, 33, 1829-1878.

Hansen, K.T., Heckman, J. and Mullen, K.J. (2004), 'The Effect of Schooling and Ability on Achievement Test Scores', *Journal of Econometrics*, 121, 39-98.

Heckman, J. (1979), 'Sample Selection Bias as a Specification Error', *Econometrica*, 47, 153-161.

Maani, S.A. and Kalb, G. (2007), 'Academic Performance, Childhood Economic Resources, and the Choice to Leave School at Age Sixteen', *Economics of Education Review*, 26, 361-374.

Maloney, T. (1999), An Analysis of Representativeness and Attrition in the CHDS, New Zealand Treasury Working Paper, September 1999.

McKenzie, K. and Schweitzer R. (2001), 'Who Succeeds at University? Factors predicting Academic performance in first Year Australian University Students', *Higher Education Research and Development*, 20, 21-33.

Micklewright, J. (1989), 'Choice at Sixteen', *Economica*, 56, 25-39.

Miller, P. and Volker, P. (1989), 'Socio-economic Influences on Educational Attainment: Evidence and Implications for the Tertiary Education Finance Debate', *Australian Journal of Statistics*, 31A, 47-70.

Montgomery, J. (1991), 'Social Networks and Labour-Market Outcomes: Toward an Economic Analysis', *American Economic Review*, 81, 1408-1418.

OECD (2008), Jobs for Youth, OECD, Paris.

Parker, J.D.A., Creque Sr., R.E., Barnhart, D.L., Harris, J.I., Majeski, S.A., Wood, L.M., Bond, B.J. and Hogan, M.J. (2004), 'Academic Achievement in High School: Does Emotional Intelligence Matter?', *Personality and Individual Differences*, 37, 1321-1330.

Paschall, M.J. and Freisthler, B. (2003), 'Does Heavy Drinking Affect Academic Performance in College? Findings from a Prospective Study of High Achievers', *Journal of Studies on Alcohol,* 64, 515-519.

Prior, H. and Beggs, J.J. (1989), 'Influence of Family Background on the Educational and Labour-Force Outcomes of Year 12 School-Leavers', *Australian Journal of Statistics*, 31A, 99-124.

Rangvid, B.S. (2010), 'Source Country Differences in Test Score Gap: Evidence from Denmark', *Education Economics*, 18, 269-295.

Rice, P.G. (1987), 'The Demand for Post-Compulsory Education in the UK and the Effects of Educational Maintenance Allowances', *Economica*, 54, 465-475.

Rothstein, J.M. (2004), 'College Performance Predictions and the SAT', *Journal of Econometrics*, 121, 297-317.

Sandy, J. and Duncan, K. (2010), Examining the Achievement test Score Gap Between Urban and Suburban Students, *Education Economics*, 18(3), 297-315.

Schweinhart, L.J., Montie, J., Xiang, Z., Barnett, W.S., Belfield, C.R. and Nores, M. (2005), Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40, Monographs of the High/Scope Educational Research Foundation, Ypsilante, Michigan, High/Scope Press.

Stinebrickner, T.R. and Stinebrickner, R. (2008), 'The Causal Effect of Studying on Academic Performance', *The B.E. Journal of Economic Analysis and Policy*, 8, article 14.

Trockel, M.T., Barnes, M.D. and Egget, D.L. (2000), 'Health-Related Variables and Academic Performance Among First-Year College Students: Implications for Sleep and other Behaviors', *Journal of American College Health*, 49, 125-131.

Willis, R.J. and Rosen, S. (1979), 'Education and Self-Selection', *Journal of Political Economy*, 87, S7-S36.