Childhood Economic Resources, Academic Performance, and the Choice to Leave School at Age Sixteen *

Sholeh A. Maani
Department of Economics
The University of Auckland

and

Guyonne Kalb
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Working Paper No. 1/03

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 0 7340 3113 0

January 2003

*We would like to thank Drs David Fergusson and John Horwood for making the data set available and for insightful information on the data set, and Adam Warner at the University of Auckland for research assistance. We are also grateful to seminar participants at the University of Wisconsin, the University of Michigan, the Institute of Social and Economic Research (ISER) at the University of Essex and the Australasian Econometric Society Meetings (ESAM 2002) for valuable comments.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 5330**
***Fax* (03) 8344 5630**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

This paper uses a recent panel data set from New Zealand to examine the link between academic performance and the decision of teenagers to leave school. These choices have significant lifetime economic impacts, since early school leaving in many cases closes pathways to further education. We address the potential endogeneity and error correlation of academic performance and later school-leaving choices. The results show that schooling decisions largely represent personal choice, and that they are influenced by factors that are at work for a long period of time. Personal ability, parental education and income during early and later childhood years influence the demand for education, exerting their influence directly and through academic performance. These results point to the role that academic performance could play in breaking cycles of disadvantage.

1 Introduction

Early school leaving in many cases closes pathways to further education. In countries such as the U.K., Australia and New Zealand, where education is compulsory up to the age of sixteen, schooling choices at sixteen have lifetime economic impacts. A general observation is that adolescents from disadvantaged families are more likely to leave school beyond post-compulsory levels, and less likely to participate in university education.[1] This paper uses a recent panel data set from New Zealand to determine the factors in the decision of school leaving and, in particular, to examine the link between academic performance in national examinations and subsequent school-leaving choices by adolescents. A simultaneous model of academic performance and school-leaving choices is estimated and error correlation between the two equations is addressed. The analysis incorporates the effects of academic ability, parental education, family resources over the child's growing years, and school and peer characteristics.

In New Zealand, education starts at age 5 with a kindergarten-equivalent year, it continues for 12 years after the initial year or for a total of 13 years, and it is compulsory up to age 16. Those who are at school in grades 10 and 12 are expected to take nationally administered examinations. These School Certificate Exams at the end of grade 10 have been in effect for decades, and are taken when students are 15 or 16 years old depending on their birth month. These exams, that are nationally administered, are based on the same set of questions and grading for all participants. This is a great advantage as the use of such a measure of academic performance eliminates problems with the potential inconsistency in comparing grades across schools in lower and higher income decile localities. The School Certificate results thus provide nationally comparable academic performance results at around age 15 and prior to potential school leaving at age 16, as opposed to, for example, SAT scores (Scholastic Aptitude Tests) for College Entrance Examinations in the U.S., which are given at the end of high school and could not be used in models of early school leaving.[2]

Demand for post-compulsory education as a personal choice has been addressed by Willis and Rosen (1979), who estimate participation in university studies, and by Rice (1987), who

[1] In addition, the experience of countries such as Australia and New Zealand, which abolished university fees to increase access, throughout the 1970s to the 1980s has shown that the socio-economic background of students is highly stable over time. For example, in Australia 10 years after the abolition of university fees, the socio-economic background of tertiary students had not changed, mainly representing children from fathers with a white-collar occupation and higher income levels (for example, Williams, 1987; Anderson and Vervoon, 1983; and Wran, 1988).

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

estimates secondary school leaving in Britain. The question of the effect of parental resources on academic performance of children and adolescents, has received recent attention by, for example, Blau (1999), Feinstein and Symons (1999), and Ermisch and Francesconi (2001). This study extends the literature by estimating models of school leaving jointly with models of academic performance and incorporates the impact of family resources and a child's personal characteristics. The study allows personal and family resources throughout childhood to influence academic performance, which in turn can influence early school-leaving choices.

The innovations of this study are the estimation approach, the addressing of the link between academic performance and early school-leaving choices of young adults, and the use of longitudinal data to explain schooling outcomes. This is made possible by the use of the Christchurch Health and Development Study (CHDS) data set[3], which allows us to control for important longitudinal personal and household conditions starting from birth, such as annual parental income decile from early childhood to teenage years, the child's IQ at age 8, and a measure of academic performance at age 15, which is standardized through the use of a national set of examinations. These features of the analysis substantially reduce the 'unexplained' part of estimations due to unobserved heterogeneity, and make it possible to control for the sets of factors that are related to personal academic ability and parental resources. The ability to make a distinction between early childhood and teenage household income conditions is of interest in examining the effects of parental income at different points in time on academic performance and early school-leaving choices.

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 of the sample follow in Section 3. The models and their results are presented in Section 4, followed by concluding remarks in Section 5.

2. Analytical Framework and Estimation

The theoretical modelling framework, which is widely adopted in the economic literature on participation in post-compulsory education, focuses on individual choice for long-term investment in human capital and the intertemporal nature of the investment decision (for

[3] Further information and other research using this data set can be found in Fergusson et al. (1989), Fergusson et al. (1991), and Fergusson and Lynskey (1993).

example, Becker, 1993; Schultz, 1961).[4]

The decision to participate in higher education and training is intrinsically related to a number of factors. For example, investment in higher education is expected to result in higher returns for those with greater academic ability and a taste for lifetime labour force participation. In addition, household financial constraints would influence the cost of obtaining education. Moreover, the family socio-economic background can affect the demand for post-compulsory and higher education through tastes and the costs of obtaining information.

Therefore, *ceteris paribus,* those individuals who have higher academic ability and a stronger taste for earned income as opposed to leisure over their lifetime are more likely to invest in higher education. Likewise, keeping ability constant, a greater potential to finance education will lead to greater participation. The model has further been extended to control for other personal characteristics such as age and gender.

An extended framework for analysing participation in higher education is based on the Willis and Rosen (1979) model of participation in university studies in the U.S. and Rice's (1987) extension to secondary school leaving in Britain. In this framework, choosing a level of education depends on the expected value of lifetime earnings at that education level and on background characteristics, which determine the individual's tastes, expectations and the financial constraints facing the household. Thus, individuals select different levels of education on the basis of financial resources, tastes, perceptions and natural ability. That is, individuals are sorted into education levels (for example, school leaving at age 16 or continuation to post-compulsory levels) according to an underlying joint distribution of tastes, talents, expectations, and parental income. These characteristics are assumed to be randomly and independently distributed across individuals. While Willis and Rosen's analysis utilized structural models and focused on self-selection, Rice's application utilized reduced form models of participation and emphasized the effect of financial constraints on school-leaving choices of males and females.[5] Neither study had observable variables on academic ability such as IQ or academic test scores.

---

[4] A detailed discussion of the human capital framework for the study of participation in higher education, and a review of the empirical literature on participation in post-compulsory and higher education is provided in Maani (1996a, 1996b, and 1997).

[5] It is interesting to note that although the Willis and Rosen (1979) model is based on Human Capital theory, it is also consistent with signalling theories of investment in education, since in both theories schooling is pursued to the point where its marginal (private) internal rate of return equals the rate of interest. Both theories are also consistent with the model in which participation in education is influenced by the capacity to finance education, ability, tastes, perceptions and information, and expectations (some observed and some unobserved) --although in human capital theory, investment in education is assumed to increase labour productivity, while in signalling theory, education is a positional good to signal information on unobserved ability.

In this framework, educational choices for the *i*th individual are influenced by

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

where V_{ij} is the utility of net expected present value of life-time earnings at each level of educational attainment j (E_j), as influenced by individual talents and abilities (S_i); observable personal and environmental characteristics (X_i), which determine the individual's tastes, expectations and the financial constraints facing the household; and u_i are the unobservable components. Thus, the individual invests in additional education beyond the compulsory level if the expected net benefits are positive ($V_{i1} – V_{i0} = G(S_i, X_i, u_i) >0$).

Estimation of the probability of enrolment at post-compulsory education (Pr PCE) is based on equation (2) below:

$$\text{Pr PCE}_i = \text{Pr}\ [\ (V_{i1} - V_{i0} = G(S_i, X_i, u_i) >0\ ] \qquad (2)$$

where vectors of observables S_i and X_i would result in participation in post-compulsory education if $V_{i1} - V_{i0}$ is positive. Assuming that the net benefits conditional on S_i and X_i are normally distributed and that G( ) is a linear function of S_i and X_i, Pr PCE_i would follow the standard normal cumulative density function and equation (2) can be estimated via probit analysis, such that

$$V_{i1} - V_{i0} \sim N\ (\ S_i'\beta + X_i'\gamma\ , \sigma^2\ ) \qquad (3)$$

with β, γ and σ^2 constant across the population (for example, Willis and Rosen, 1979; Rice, 1987).[6]

Academic performance, as measured by the average score in a set of examinations, can in turn be modelled as:

$$A_i = f(S_i, X_i, v_i) \qquad (4)$$

where A_i is the score of individual i (representing ability, effort, and parental investments); S_i represents personal talents and abilities; X_i is a vector of personal and parental resources, and environment; and v_i represents the effect of unobserved factors, such as motivation or

[6] The above model is nested in a model of lifetime utility maximization, which determines labour supply and education investment decisions. Although it is possible to estimate empirical models, which are based on joint determination of expected future labour supply and participation in higher education if sufficient data is available, the education participation model above presents a satisfactory approach by providing a reduced-form model of participation, which incorporates the effect of tastes and ability. In addition, the life-time supply decisions of young persons have not materialized at the time of participation in education, and they can at best be measured empirically as stated expectations or intended behaviour influenced by the same set of factors which determine the participation in education decisions.

traumatic events prior to exams.

Academic performance of children and adolescents has recently received interest in the literature in models that link parental resources to children's academic performance. Examples of studies on the effect of parental resources on educational attainment or labour market outcomes for their children are Blau (1999), Borjas (1995), Case and Katz (1991), Duncan et al. (1998) and Montgomery (1991) for the U.S.; Feinstein and Symons (1999) and Ermisch and Francesconi (2001) for the U.K.; and Miller and Volker (1989), Prior and Beggs (1989), and Borland and Wilkins (1996) for Australia.

The recent literature on educational attainment has emphasized the significance of parental investments in human capital since childhood. An important implication of this literature is the recognition that even if teenagers have a significant input into the decision of when to terminate secondary schooling, they are constrained in their choice by academic performance and human capital investments throughout childhood (Ermisch and Francesconi, 2001).[7]

While the models of *children's* academic performance are usually based on a production function, where the parents are the producers, the academic performance of teenagers at age 15 and their school-leaving choices at age 16 are likely to be joint decisions, influenced, amongst other things, by the adolescent's personal ability and human capital investments by parents throughout childhood.

If the score A_i is a continuous measure, equation 4 can be estimated through Ordinary Least Squares. In this case a Tobit specification is used, because of the censored nature of exam grades below the Fail (grade D) and above the grade A cut offs. That is, the academic performance of the most capable and the least capable students cannot be accurately measured through an exam targeted at the average student.

In the econometric discussion of the modelling approach in Section 4, we consider joint estimation issues of equations 4 and 2, and deal with correlation across academic performance and school-leaving choices of individuals. In joint estimations, academic performance is included in the school-leaving choice and vice versa, because academic performance is likely to be influenced by prior school-leaving intentions and vice versa academic performance can influence the decision to drop out of school. In addition, the same set of unobserved variables can potentially influence both academic performance and school leaving, which means the error terms in the two equations may be correlated.

The study further extends earlier studies by examining the effect of personal ability and

---

[7] Ermisch and Francesconi (2001) consider seven levels of educational attainment in the U.K., from no qualifications to degree qualifications, and estimate ordered logit models of educational attainment. Other studies such as Feinstein and Symons (1999) have focused on test scores in mathematics and English at age 16

parental resources at different points in time on both academic performance and school-termination choices of young adults, using a data set, which includes a remarkably large number of relevant variables, thereby reducing the importance of the unobserved components. Nevertheless, the joint estimation of the equations for 'academic performance' and 'dropout' in the subsequent year allows us to formally account for the joint decision on school leaving and academic performance, and to allow for the potential correlation of the unobserved components.

3. Characteristics of the Sample

The Christchurch Health and Development longitudinal Study (CHDS) includes extensive economic and academic information on a cohort born in Christchurch in 1977. This cohort is followed throughout their childhood and adolescence, providing information on their transition from school to further education, training and work. 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 this study utilizes information from survey years from birth in 1977 to age 16 of the cohort, selecting respondents for whom data on all variables of interest was available.[8] The characteristics of the sample are summarized in Table 1 below.

Table 1 presents the mean characteristics of the sample. Columns 1 and 2 identify mean characteristics of individuals in the sample and those who continued with post-compulsory education. These characteristics include the individual's IQ at age 8; the average School Certificate grade obtained (reflecting academic performance); the household income decile between ages 11 and 14 and in early childhood between ages 1 and 5; and school, neighbourhood and peer factors, such as the proportion of the student's class continuing to post-compulsory levels (at age 16) or association with peer groups with deviant behaviour (a 1-10 scale reflecting problems with the law, substance abuse, etc.), which is expected to serve as a proxy for the student's interest in deviant behaviour.

---

in the U.K.

[8] The original cohort of individuals in the survey consisted of 1265 individuals. The sample used in this study contains 713 observations to analyse the dropout before exams and 598 observations for the joint estimations of School Certificate Examination at age 15 and school leaving at age 16. The smaller sample used for age 16 is partly due to minor attrition over time, and partly due to missing values on variables of importance to this part of the study, such as average score, parental income, and school factors. Analysis indicates that the selected sample is slightly less likely to drop out of secondary school than the full sample (the probability is 0.0034 lower). Maloney (1999) showed that attrition was related to some initial characteristics such as ethnicity and having a single parent. Nevertheless, comparisons with later Census data at both local national levels show that the CHDS is still fairly representative of the population of children born around 1977.

Table 1: Characteristics of the Sample

Characteristics	**Mean** (Standard Deviation) Full Sample	Continued with Post-compulsory Education
Personal Characteristics		
Female (%)	50.5%	52.6%
Maori Ethnicity (%)	7.4%	5.7%
Pacific Island Ethnicity (%)	2.8%	2.3%
IQ (tested at 8 years of age)	102.8 (14.88)	105.8 (13.63)
Education		
Average School Certificate Grade Point Average (age 15 and 10th grade, where Fail=0, C=1, B=2, A=3)	1.061 (0.86)	1.26 (0.80)
Mother without Qualifications (<10th grade)	49.8%	43.6%
Mother with a Tertiary Qualification	20.6%	24.6%
Father without Qualifications (< 10th grade)	47.5%	41.3%
Father with a Tertiary Qualification	19.8%	23.2%
Total Dropout rate from school at Age 16	15.5%	---
Dropout rate from school before Exams	8.8%	---
Family and Social Environment		
Adolescent Average Income Decile: Ages 11-14 (10 is most affluent decile)	5.53 (2.54)	5.92 (2.48)
Early Childhood Average Income Decile: Ages 1-5	5.82 (2.40)	6.13 (2.33)
Own their Home (%)	88.6%	92.1%
Number of Siblings	1.48 (0.94)	1.47 (0.89)
Rural Location (%)	15.9%	16.0%
Percentage of Family Income from Benefits	13.9%	10.6%
Regional Unemployment Rate	10.6%	10.6%
Percentage of Respondent's class continue at Age 16	83.6% (16.2)	86.4 (11.6)
Average Class Size	28.8 (4.20)	28.9 (4.15)
Association with Deviant Peers age 15 (10 is the highest association)	2.30 (2.45)	1.91 (2.14)
Sample Size:	**713**	**561**

The IQ variable is the total IQ score at age 8 based on the Wechsler Intelligence Scale for Children (WISC) testing the cognitive performance of children, which is an international test performed by qualified psychologists only. The scores range from 70 to 145. We expect that childhood cognitive tests reflect both innate ability and investments by parents in their children up to that time, and it is useful to have this information in the models to predict future academic performance.[9]

As column 1 of Table 1 on the full sample shows, about half of the sample (50.5 per cent) was female. The characteristics of the sample on academic performance and economic conditions are reassuring in relation to expected national averages, such as the average IQ of 102.8, and the average school certificate mark of 1.06 or a C, which is the average national grade for these exams. 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 were Maori and 2.8 per cent were Pacific Islanders.

In relation to parental education, 49.8 per cent of the mothers and 47.5 per cent of the fathers of the respondents had no school qualifications (that is, less than the year 10 School Certificate), and 20.6 per cent of mothers and 19.5 per cent of fathers had tertiary (university or other post high school) qualifications.

In general, the mean group characteristic comparisons show that those who continued in post-compulsory education had mean characteristics, which were different from the full sample. These differences included a higher average IQ at age 8, higher average School Certificate marks, they belonged to a higher family income decile, their parent's education level was higher and they went to a school with a higher proportion of the class continuing to the post-compulsory 11th grade (or Sixth Form). These characteristics are consistent with the hypothesis that individuals sort themselves into different choices based on their academic ability and the expected returns to their choice, family income constraints (lower income deciles are more likely to leave school early and less likely to participate in higher studies), and influences from the school and peer environment.

4. Models and Results

In this section, academic performance and the early school leaving of adolescents is discussed with an emphasis on the role of academic performance in this process.

---

9 This IQ variable is correlated with both reading and mathematics scores at ages 8 and 9, but we found that the IQ score was consistently a better predictor of later academic performance than the other two measures, which possibly reflects the measurement of a broader set of skills.

Section 4.1 describes the basic underlying model of interest of school leaving at 16 years of age, in which the potential endogeneity of academic performance is explored. Section 4.2 describes the specific model that is estimated and presents the results. Several sensitivity analyses are discussed in Section 4.3.

4.1 School Leaving at Age Sixteen

The school-leaving model examines the effect of personal characteristics such as cognitive ability and academic performance, peer and school effects, and parental economic constraints on the choice to drop out of school at age 16. The dependent variable is binary as to whether or not the respondent had left school at the post-compulsory age of 16, as opposed to enrolment in the 11th grade (also called the Sixth Form), and beyond the School Certificate. A probit model is therefore chosen to analyse school leaving:

$$\Phi^{-1}(\Pr PCE_i) = \alpha + A_i'\beta + X_i'\gamma \qquad (5)$$

where Φ^{-1} is the inverse of the standard normal cumulative distribution function, $\Pr\ PCE_i$ is the probability that the respondent had left school at age 16, A_i represents academic performance for individual *i,* and X_i represents personal characteristics such as ability, gender, socio-economic and cultural background, and household and environmental constraints such as household assets, proportion of household income from government benefits, school effects, and neighbourhood and peer effects .[10]

4.1.1 Academic Performance and its Potential Endogeneity

Academic performance is measured by the average score in the National School Certificate Examination, which is taken at the end of the 10th grade at age 15 or 16, on 5 subjects.[11] If a student fails this examination they would need to repeat the year. 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.[12] The average score is constructed by averaging the numeric values of the score over all subjects taken in the certificate. Thus the minimum score is 0 for a D (fail) and the maximum score is 3 for an A average and as a result the average grade on all subjects is censored at 0 and 3. Individuals below and above a certain academic performance level cannot be ranked besides observing that they are at the minimum or the maximum level. Truncation is most prevalent at the lower end in this data set.

---

[10] See Table A1 in Appendix A for a full list of variable descriptions.

[11] With new terminology implemented in 2002 in New Zealand, the years of study from age 5 to 18 are termed Year 1 to Year 13, but in this paper, we are referring to the North-American equivalent years of kindergarten to the 12th grade.

[12] This is also compatible with the official GPA (Grade Point Average) score assignment in New Zealand.

Academic performance may be endogenous as it is a combination of ability and effort. Academic performance at the end of the 10^{th} grade is expected to be influenced by many personal, school and family resource variables, which also influence school-leaving choices at age 16. In addition, academic performance is likely to be influenced by school-leaving intentions, and academic performance can influence the decision to leave school. Finally, a similar set of unobserved variables can potentially influence both academic performance and school leaving. In recognition of the simultaneous effects and potential error correlation across the two equations, we pursue joint estimation of academic performance and the school-leaving choice, even though the data set used allows us to control for a remarkably large number of important variables, which are often not observed in other data sets and are likely to reduce the error component of the two equations.

The equation for latent academic performance A_i* looks as follows:

$$A_i^* = \alpha_a + X_{ai}'\gamma_a + \varepsilon_a \qquad (6)$$

However, we only observe A_i, the average score, which is censored at the lower and upper end:

$$A_i = A_i^* \text{ if } 0 < A_i^* < 3 \qquad (7)$$

$$A_i = 0 \text{ if } A_i^* \le 0 \qquad (8)$$

$$A_i = 3 \text{ if } A_i^* \ge 3 \qquad (9)$$

This equation is estimated using a standard Tobit specification. The joint model including direct effects of the endogenous variables in both equations looks as follows:

$$A_i^* = \alpha_a + X_{ai}'\gamma_a + PCE_i\delta_a + \varepsilon_a \qquad (10)$$

$$D_i^* = \alpha_d + X_{di}'\gamma_d + A_i\delta_d + \varepsilon_d \qquad (11)$$

Where D_i^* is the latent dropout.

In the joint model, the academic performance equation is estimated by Tobit and the dropout equation by Probit, as before. The observed endogenous variables are used on the right hand side, because it seems unlikely that the effect of having an extremely good or bad score will be different from having a very good or bad score. Similarly, the actual decision on school leaving (PCE) is likely to be the important variable rather than the latent dropout, that is, being extremely unlikely to drop out of school will probably not have a different effect on academic performance compared to being very unlikely to drop out.

4.1.2 Other variables

The measures of parental income we have included reflect permanent rather than current income providing a measure of long-term parental resources (see Blau, 1999, for U.S. evidence based on the NLSY data set). In addition, the data allow us to distinguish between parental income effects during early childhood and adolescent years on later academic performance and school-leaving choices. We have chosen average income deciles over two distinct time periods with a gap between them of early childhood and adolescent years, to explore the effect of income in two distinct time periods. The 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. 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. This distinction between income at different points in time may be relevant as is shown, for example, by Duncan et al. (1998), who find evidence for the U.S. (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.

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 significance of benefit income compared to the young person's family income. The variable also reflects beneficiary status and relative disadvantage with regard to the household's wealth and assets.

4.2 Academic Performance and School Leaving (Joint Estimation Results)

The results from the joint estimation of the dropout and the academic performance equations in Table 2 provide strong evidence that school-leaving choices of young adults are influenced by a host of factors, among which academic performance has a prominent place. Academic performance, as shown, is in turn influenced by parental education and resources, personal ability, and school and peer effects. These are all factors that are at work over a number of years before the decision on school leaving at age 16. The academic performance equation in Table 2 also shows a different effect of the intention to drop out on the average score in examinations.

The effect of the joint estimation compared to separate estimations, including the same variables, is that in the joint estimation the direct effect of academic performance on school-leaving choices becomes more dominant (its coefficient increases from –0.710 to –1.180), and the effect of school leaving on academic performance becomes smaller (the coefficient

Table 2: Academic Performance and School Leaving
Joint Estimation:
(Equation (10): National Exam Grade at age 15)
(Equation (11): Dropout: 1=Left School at Age 16; 0=Enrolled in School at Age 16)

Explanatory Variables	Average_Grade (TOBIT)			Dropout (PROBIT)		
	Coefficient	z-value	$P>\|t\|$	Coefficient	z-value	$P>\|z\|$
Direct Effects						
Ave_Grade				-1.1805	-4.57	0.000
Dropout	-.2982	-2.43	0.015			
Personal Characteristics						
Female	.1883	3.40	0.001	-.4736	-1.78	0.075
Maori	.1466	1.34	0.181	.3898	1.27	0.203
P_Island	-.1111	-0.65	0.515	-.3271	-0.66	0.512
IQ8	.0280	13.01	0.000			
Family & Social Environment						
M_No_Q	-.1317	-2.08	0.037	.0799	0.34	0.733
M_Tert_Q	.1164	1.58	0.115	-.0223	-0.06	0.950
F_No_Q	-.0256	-0.42	0.677	.0713	0.33	0.743
F_Tert_Q	.1371	1.76	0.078	.0493	0.14	0.888
Inc_Decile (ages 11-14)	.0567	3.83	0.000	.0361	0.59	0.554
Early_Inc_Decile(ages1-5)	.0409	2.93	0.003	.0986	2.07	0.038
Num_Siblings	.0145	0.48	0.633	.2270	2.12	0.034
Own_Home	-.1644	-1.58	0.114	-.3669	-1.26	0.209
Rural	.0437	0.60	0.547	.7954	2.15	0.031
Ben_Prop	.2544	2.26	0.024	.7436	2.23	0.026
Local_Unemp				.4132	1.16	0.246
Prop_Continue	.4825	1.99	0.047	-.2231	-0.27	0.790
Class Size				.0302	1.14	0.254
Peer_Dev	-.0554	-4.52	0.000	.0756	1.92	0.054
Constant	-2.5942	-7.95	0.000	-6.5258	-1.60	0.109
_sigma (Ancillary parameter)				.6278	31.92	0.000
Rho (correlation coefficient)				.4138	2.08	0.037

Observation summary:

Number of observations = 598 — LR $chi^2(17)$ = 402.00

Log likelihood = -674.4762 — Prob > chi^2 = 0.0000

decreases from -0.407 to –0.298). Thus, separate estimation underestimates the direct effect of academic performance on dropout and overestimates the direct effect of dropout intentions on academic performance.[13]

The results of the equation for Academic performance, show that both early childhood and adolescent parental income deciles are important in explaining academic performance. Both recent and early childhood income decile variables are highly significant. 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 0.056 of a grade in the exams. In addition to this effect, early childhood income decile explains an additional effect of about 0.04 of a grade. Therefore, keeping other factors constant, together the predicted effect of the income decile variables is close to 0.1 of a grade difference for each income decile or close to a complete grade for 10 deciles difference (the difference between a C or a D average grade, for example). The results found here are consistent with the effect of poverty (Duncan et al., 1998); the positive effect of income in many U.S. studies (Haveman and Wolfe, 1995); 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 and high school completion, is not repeated here for academic performance. These results are further consistent with recent results for the U.K. (Feinstein and Symons (1999) and Ermisch and Francesconi (2001)) regarding the importance of resources throughout childhood in determining children's academic performance.

As the results in Table 2 show, academic ability is also important. Each additional unit of IQ score, was in turn, associated with 0.028 of a grade. 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 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 is equivalent to slightly more than 3 IQ units.

In addition, the mother's lack of school qualifications, class and peer effects are significant in explaining academic performance, and females perform better academically.

Results in the dropout equation in Table 2 show that, keeping other factors constant, females are less likely to drop out of school, those who had more siblings were more likely to drop out, and early childhood income has an additional and significant effect on school leaving,

---

[13] Results based on the separate estimation of equations 10 and 11 are provided in Tables A2 and A3 in the Appendix.

even after controlling for academic performance. The effect of private versus public primary and secondary schooling was also estimated and later eliminated due to consistent insignificance.

There is some evidence that the beneficiary status of the family, as measured by the proportion of gross income from government benefits, was associated with the probability of school leaving, even after controlling for factors such as academic performance. This variable, which measures the extent of beneficiary status, is expected to reflect relative disadvantage in terms of parental assets, relative income, and other disadvantage in terms of information or social networks.[14]

It is interesting to note that once personal, socio-economic and environmental characteristics are controlled for, Maori and Pacific Island teenagers do not perform more poorly and they do not have a significantly higher probability of dropping out of school. In the dropout equation, once academic performance and peer and school effects are included in the dropout model, relatively few other variables are significant. All variables for ethnic background and indicators for parental education were insignificant.

Similar to the finding by Duncan et al. (1998), early income deciles have a greater effect on school leaving than later income deciles. With the inclusion of variables on academic performance, the proportion of family income from benefits, and school and peer effects from age 15, the income decile for adolescent years does not represent an additional direct effect on school leaving. However, early childhood income does show a direct effect on school leaving, by providing additional information on family resources over time, not contained in the later income variables.

The correlation between the error terms in the two equations is significantly positive. This is consistent with the error terms containing unobserved 'traumatic' effects. For example, students who had performed unexpectedly poorly in their exams due to an exogenous (and unobserved) incident such as an accident, illness, or emotional trauma before the exams, but continued with their post-compulsory education due to long-term choices, would have had lower than expected exam marks (negative error terms in equation 10). However, they also would have had a lower than expected dropout rate according to the model given their low exam score (negative error terms in equation 11).

Usually one would expect the effects of unobserved ability or motivation and effort to

[14] 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. Since the benefit ratio in this study is the only measure of current income (that is, household income when the child is aged 16), the negative effect on school retention probably reflects the effect of economic disadvantage.

introduce a dominant negative component to the error correlation, because similar unobserved factors would have a positive effect on the average score and a negative effect on the dropout probability, causing a negative correlation of the error terms. Therefore, one would expect the correlation coefficient to become negative if important variables are excluded from the equations.

We further investigated the sign of the error correlation (ρ) by excluding the direct effects of 'dropout' and 'academic performance' in the joint estimations, while allowing for non-zero error correlation. We found that this causes ρ to become significantly negative, which is the sign we would expect if the error terms contained important unobserved characteristics like motivation, as opposed to unobserved (traumatic) events. We suspect that average score may contain information on motivation and effort besides the obvious information on ability, so that the error term in the dropout equation no longer contains these variables.

4.3 Sensitivity Analysis

In this subsection, we compare the results from Section 4.2 with some alternative specifications.

4.3.1 Test for sample selection

The relation between the probability of dropout and average score is further explored by examining the average score and dropout before the exam. The results in the previous section are only based on those who had not left school before the exam. Since we can only observe the average score for people who did not drop out before the exam and who took the exams, we need to allow for sample selection in the observed exam results. For the students who had left school before the exams, no measure of academic performance is available, but fortunately information on all other factors is available. This means we can estimate the probability of having observable examination scores (by not having left school prior to exams) with a probit model

$$\Phi^{-1}(\Pr PBE_i) = Z_i'\gamma_Z \qquad (12)$$

where $\Pr PBE_i$ is the probability of dropping out before the exams and Z_i are the explanatory variables and γ_z are the parameters to be estimated.

Since academic performance is estimated by a Tobit, the usual Heckman adjustment for sample selection does not apply. However, we can estimate equations (12) and (6) jointly through maximum likelihood and test for their independence. In the sample selection probit, a dummy variable on birth month (=1 if birth was during the first semester of the year) was included for identification. Birth month can be used for identification because 5-year olds who start school in the first semester take the first kindergarten equivalent year only once, but those who are born in the second semester usually have a full kindergarten year in the following year. They could therefore reach age 16 and leave school before the national examinations at the end of the 10th grade. The results showed that those who were born in the second semester were indeed more likely to have left before the exams.

The error correlation coefficient across the two equations in the joint estimation failed the test of significance (estimated at –0.138 with a z-value of -0.42). In addition, joint estimation resulted in minor changes in the other parameters and the log likelihood value.[15] Different specifications, that is including or excluding birth month, gave consistent results of the independence of the two equations, indicating that the two equations were independent.[16]

Given these results, which indicate that sample selection is not an issue, we estimate equation (6) for everyone who had examination results without correcting for sample selection.[17] Using the estimated parameters from the academic performance Tobit to predict the average grade for everyone in the sample, we find that those who had dropped out before the national exams were expected to perform poorly and significantly below those who did take the national exams (see Table 3). The predictions in the first row, for example, show that those who took the national exams had a predicted average grade of C (mean =1.1599) compared to a predicted grade of D (mean=0.2635) for those who did not. This expectation of poor performance is due to a difference in observed characteristics between the two groups and unobserved characteristics do not seem to play a role.

4.3.2 Sensitivity analysis to alternative specifications

The identifying variables and exclusion restrictions we have chosen for equations (10) and (11) are IQ at age 8, which is consistently a significant determinant of academic performance at age 15, but does not significantly influence the dropout probability when academic performance at age 15 is present in the model. In the separately estimated dropout equation, IQ is also insignificant when the average score is included. This can be explained by the fact that academic performance represents a combination of academic ability (as could be measured by IQ scores), effort and parental investments. It thus contains more information than IQ alone and should include all information that is contained in IQ. In equation 11, we have included two extra variables: regional unemployment rate by gender for 15 to 19 year olds, which is expected to influence job opportunities and thus school-leaving choices, and

[15] In order to test for the robustness of this result further, we also used the usual two-step Heckman adjustment and estimated equation (6) through OLS rather than Tobit, given that many observations on academic performance are not censored. The earlier results were fully confirmed by this exercise.

[16] These results are available from the authors.

[17] The Tobit had a number of significant coefficients, especially the income variables, IQ, gender, mother's qualifications, and peer variables (the results are available from the authors).

average class size throughout secondary school, which is consistently insignificant in the academic performance but not in the school-leaving equations.

Table 3: Predicted Average Grade for Full Sample, based on Tobit Regressions

	Mean	Standard Deviation	Minimum	Maximum
Group 1: Did Not drop-out before, and took exams (n=629):	1.1599	.6151	0	3
Group 2: Did Not dropout before Exams, but Did Not take exams (n=21):	.3453	.3796	0	1.5379
Group 3: Dropped out before exams, and Did Not take exams (n=60)	.2635	.3782	0	1.3986

Note: Estimations for the full sample, based on estimates for all those who took the national level School Certificate Exams at the end of the 10^{th} grade.

The validity of the identification strategy is investigated by including all variables in both equations. This is possible, because the model is in principle identified through functional form. The results from this estimation show that IQ is highly insignificant in the dropout equation, indicating that after taking into account the average exam score, the IQ score is not relevant anymore (see Table 4). The coefficients for the regional unemployment rate and for the average class size in the average score equation are also insignificant and slightly positive.

The estimated effects of the other variables do not change much after including these additional variables. In particular, the effect of the average score variable and the correlation variable remain similar to the effect estimated here, showing the robustness of the effects of interest to this change in specification. Testing between inclusion and exclusion of the three variables discussed above, we find a Chi-square of 4.18 with three degrees of freedom. This means we cannot reject the model, which excludes the extra variables, at the 5-per cent level.

The robustness of the results was further investigated by examining variations around the choice of exclusion restrictions. First we excluded the number of siblings in addition to average class size and local unemployment in the average score equation and then we only excluded the number of siblings from the average score equation. It was found that the coefficients of interest are quite robust across these specifications.

Table 4: Selected Coefficients in Alternative Specifications of the Joint Model
Equation (10): Academic Performance
Equation (11): Dropout

	Coefficients in Alternative Model Specifications:				
Exclusions:	**As in Table 2:**			**None**	**Direct Effects**
In Eq. (10):	**U,C**	**S**	**U,C**		**D**
Eq. (11):	**IQ**	**IQ**			**D**
Equation 10:					
Direct Effect: Dropout	-.298*	-.307*	-.290*	-.301*	--
Inc_Decile	.057*	.057*	.056*	.057*	.063*
Early_Inc_Decile	.041*	.041*	.041*	.042*	.039*
IQ8	.028*	.028*	.028*	.028*	.030*
Equation 11:					
Direct Effect: Ave_Grade	-1.180*	-1.170*	-1.228*	-1.236*	--
Inc_Decile	.036	.035	.039	.039	-.029
Early_Inc_Decile	.098*	.097*	.099*	.100*	.058
IQ8	--	--	.002	.002	-.025*
Rho : (error correlation coefficient)	.414*	.406*	.450	.460	-.223*
Log likelihood	-674.48	-672.65	-674.47	-672.39	-683.04

Notes: Exclusions: IQ (IQ8), U (Local Unemployment), C (Class Size), S (Number of Siblings), and D (Direct Effect of Academic Performance and Dropout); *: Significant at the 5-% level.

The error correlation coefficient (ρ) between equations 10 and 11 is consistently around 0.4 in all these specifications. Except where, for example, IQ is included in the dropout equation. This inclusion renders the correlation coefficient insignificant (but at around the same value), indicating that despite its insignificance, inclusion of the IQ variable affects the error term and changes error correlation across the two equations. [18] A possible explanation for this may be that in cases of traumatic events, IQ rather than the error term picks up some of the effect of ability which makes people stay in school even if their scores are disappointing because of external events.

4.3.3 Reduced-Form Analysis of Academic Performance and School Leaving at Age Sixteen

Joint estimation has allowed us to analyse the effect of academic performance on the school-leaving choice and allow for the endogeneity of academic performance. We are interested in the comparison of the results from the joint model with the simpler reduced forms. We use the latent endogenous variables on the right hand side following the literature, so results can be more easily compared:[19]

$$A_i^* = \alpha_a + X_{ai}'\gamma_a + D_i^*\delta_a + \varepsilon_a \qquad (13)$$

$$D_i^* = \alpha_d + X_{di}'\gamma_d + A_i^*\delta_d + \varepsilon_d \qquad (14)$$

The reduced form model for academic performance, which is obtained by substituting equation (14) into equation (13):

$$A_i^* = \alpha_r + X_{ri}'\gamma_r + \varepsilon_i \qquad (15)$$

Where X_{ri} contains all variables in X_{ai} and X_{di}, and likewise, the reduced form for school leaving is obtained by substituting equation (13) into equation (14):

$$D_i^* = b_r + X_{ri}'d_r + \varepsilon_i \qquad (16)$$

Equations (15) and (16) are straightforward to estimate and the results can easily be compared to other studies of academic performance and educational choice.

Comparing the 'reduced form' and 'joint estimation' results for academic performance, the reduced form model performs well and most important results remain similar (see Table 5). For example, the combined effect of the coefficients for early childhood and adolescent income is respectively 0.10 and 0.097 in the two estimations, and the effect of childhood IQ is respectively 0.030 and 0.028.

---

[18] Other functional specifications of the model were examined as well: a bivariate probit model of either 'passing' or 'failing' the exams and subsequent school leaving; and joint estimation of an ordered probit (using three levels, fail, maybe a pass and pass) and a binary probit, using the exam results and school leaving. We found that the main results are not sensitive to these specifications, that is, academic performance is important in the drop out equation and the direction of the effects of other variables are very similar. We proceeded with the Tobit as it provides the most detailed information on the exam results.

[19] In addition, using reduced form equations with the observed endogenous variables is impossible in this

Table 5: Reduced-Form Estimations

Equation (15): Academic Performance; Equation (16): Dropout

Explanatory Variables	**Average_Grade (TOBIT)**			**Dropout (PROBIT)**		
	Coefficient	**z-value**	**P>|z|**	**Coefficient**	**z-value**	**P>|z|**
Personal Characteristics						
Female	.1844	2.97	0.003	-.7106	-2.96	0.003
Maori	.1054	0.94	0.345	.3026	0.94	0.348
P_Island	-.1099	-0.56	0.575	-.2010	-0.40	0.692
IQ8	.0298	14.12	0.000	-.026	-3.13	0.002
Family & Social Environment						
M_No_Q	-.1403	-2.18	0.030	.2168	0.95	0.343
M_Tert_Q	.1172	1.57	0.118	-.2379	-0.67	0.504
F_NO_Q	-.0375	-0.60	0.550	.0990	0.44	0.658
F_Tert_Q	.1351	1.71	0.087	-.1409	-0.40	0.688
Inc_Decile (ages 11-14)	.0610	4.06	0.000	.0326	-0.55	0.583
Early_Inc_Decile (ages 1-5)	.0392	2.75	0.006	.0640	1.29	0.196
Num_Siblings	.0056	0.18	0.855	.2217	2.02	0.044
Own_Home	-.1197	-1.12	0.263	-.3608	-1.25	0.210
Rural	.1610	1.50	0.135	.7434	2.15	0.032
Ben_Prop	.2008	1.78	0.075	.3163	1.23	0.220
Local_Unemployment	.1413	1.42	0.157	.4786	1.42	0.156
Prop_Continue	.6089	2.46	0.014	-.8825	-1.06	0.288
Class_size	.0063	0.91	0.365	.0275	1.01	0.312
Peer_Dev	-.0627	-5.13	0.000	.1183	3.10	0.002
Constant	-4.6537	-4.05	0.000	-4.5215	-1.15	0.251
Variance of the error term:	0.6325	Std. Error=.0198				
	Log likelihood = -578.94			Log likelihood = -106.55		
	LR $chi^2(18)$ = 395.06			LR $chi^2(18)$ = 64.47		
	Prob > chi^2 = 0.0000			Prob > chi^2 = 0.0000		
N=598	Pseudo R^2 = 0.2544			Pseudo R^2 = 0.2323		

The reduced form school-leaving model, however, does not perform as well. Most importantly, by having omitted the information contained in academic performance, the reduced form model reflects the effect of childhood IQ, but it cannot show the full effect of family resources on schooling choices or their effect through the impact of academic

specification.

performance. The advantage of joint estimation is that the direct effect of the academic performance can be shown explicitly. Thus, the reduced form results show the usefulness of the reduced form approach for estimating academic performance models, but they highlight the need for the joint estimation approach adopted for answering the questions on school leaving addressed in this paper.

4.3.4 Summary of Sensitivity Analyses

Overall, the results are consistently showing that academic performance is an important determinant, which sorts young persons into early school leavers and those who continue in further education. The results are compatible with the hypothesis that young persons from lower socio-economic backgrounds are influenced throughout their schooling years by parental inputs, school and peer effects such that by age 16 their academic performance, tastes and information about alternative opportunities are more likely to sort them into the school-leaving choice.

5. Conclusion

This study has provided empirical evidence on the effect of personal characteristics and family resources on the academic performance of young persons and their choices about leaving secondary school or continuing with post-compulsory education at age 16. This is an important decision since it generally determines further higher education choices and has lifetime income effects.

The individual level and longitudinal nature of the Christchurch Health and Development data set have allowed modelling and testing of a number of relevant factors. In particular, the analysis incorporated the effect of academic ability (as measured by IQ score at age eight), academic performance (as measured by an average of 10th grade national exam scores at age 15), school and peer effects, and household economic conditions over time. The study extends the literature by considering the potential endogeneity of academic performance in the school-leaving choice and by addressing the correlation between the unobserved terms in the academic performance and school–leaving equations. The results provide support for adopting the joint estimation approach taken and are robust over alternative specifications.

The analysis of the school-leaving choice at age 16 indicates that this decision is influenced by factors that are at work for a long period of time. Personal ability, household income constraints and socio-economic background are all influential in school-retention choices and exert an influence through factors such as academic performance, parental income, and school and peer effects. Not only is academic performance at age 15 influenced by parental income and resources, as reflected by the recent income decile, but early childhood income resources also have an additional and prominent impact on academic performance. The estimated coefficient of the latter is about two thirds of the coefficient size of the more current income decile variable. This result, in addition to the effect of the mother's school qualifications and childhood IQ highlights the long-term nature of the early school-leaving decision process, the outcome of which is observed at age 16. This result is consistent with recent results for the U.K. regarding the importance of early childhood resources in determining children's academic performance (Feinstein and Symons, 1999; Ermisch and Francesconi, 2001) and the direct effect of early childhood resources on the probability of school leaving (Duncan et al., 1998). This points to the need for long-term solutions to improve school retention of children from disadvantaged families.

The results are further consistent with the hypothesis that students sort themselves into schooling choices at age 16, based on the expected returns to these choices, their tastes, and information available to them through their family, school and peer networks. In the school-leaving choice, the student's academic performance is an important channel through which personal ability and economic factors exert their influence. To the extent that the academic performance of students can be influenced throughout their education years, these results point to the role that academic performance can play in breaking cycles of disadvantage.

References

Anderson, DS and Vervoon, AE (1983), *Access to Privilege: Patterns of Participation in Australian Post-secondary Education,* Canberra: Australian National University Press.

Becker, GS (1993), *Human Capital,* third edition, Chicago University Press.

Blau, DM (1999), "The Effect of Income on Child Development", *The Review of Economics and Statistics,* 81 (2), 261-276.

Borjas, GJ (1995), "Ethnicity, Neighbourhoods, and Human-Capital Externalities", *American Economic Review*, 85(3), 365-390.

Borland, J and Wilkins, R (1996), "Earnings Inequality in Australia", *Economic Record*, 72(216), 7-23.

Case, AC and Katz, LF (1991), "The Company You Keep: The Effects of Family and Neighbourhood on Disadvantaged Youths", NBER Working Paper 3705.

Duncan, GJ, Yeung, WJ, Brooks-Gunn, J and Smith, J (1998), "How Much Does Child Poverty Affect the Life Chances of Children?", *American Sociological Review*, 63, 406-23.

Ermisch, J and Francesconi, M (2001), "Family Matters: Impact of Family Background on Educational Attainment", *Economica,* May*,* 137-156.

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

Fergusson, DM, Horwood, LJ and Lloyd, M (1991), "Confirmatory Factor Models of Attention Deficit and Conduct Disorder", *Journal of Child Psychology and Psychiatry*, 32(2), 257-274.

Fergusson, DM, Horwood, LJ, Shannon, FT and Lawton, JM (1989), "The Christchurch Child Development Study: A Review of Epidemiological Findings", *Paediatric and Perinatal Epidemiology*, 3, 302-325.

Fergusson, DM and Lynskey, MT (1993), "Maternal Age and Cognitive and Behavioural Outcomes in Middle Childhood", *Paediatric and Perinatal Epidemiology*, 7, 77-91.

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.

Haveman, R and Wolfe, B (1995), "The Determinants of Children's Attainment: A Review of Methods and Findings", *Journal of Economic Literature,* 33, 1829-78.

Maani, SA (1996a), "The Effect of Fees on Participation in Higher Education: A Survey of

OECD Countries", *New Zealand Economic Papers*, 30 (1), 55-86.

Maani, SA (1996b), *A Research Agenda, Methodologies, Models and Data Requirements for Estimating Participation in Post-compulsory Education in New Zealand,* Report prepared for the New Zealand Treasury, Wellington.

Maani, SA (1997), *Investing in Minds: The Economics of Higher Education in New Zealand,* Institute of Policy Studies, Wellington.

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

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(5), 1408-1418.

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

Rice, PG (1987), "The Demand for Post-Compulsory Education in the U.K. and the Effects of Educational Maintenance Allowances", *Economica*, 54 465-475.

Schultz, TW (1961), "Investment in Human Capital", *American Economic Review, 51, 1-17.*

Williams, T (1987), *Participation in Education,* Research Monograph No. 30, Hawthorn, Victoria: Australian Council for Educational Research.

Willis, RJ and Rosen, S (1979) "Education and Self-Selection", *Journal of Political Economy*, 87, S7-S38.

Wran, N (Chairman) (1988), *Report of the Committee on Higher Education Funding,* Department of Employment, Education and Training, Canberra: Australian Government Printing Services (the Wran Report).

APPENDIX A
Table A1: Definition of Variables

Dropout	Binary dependent variable: 1 for individuals who did not enrol beyond the post-compulsory level (age 16, usually at the 11th grade (Sixth Form); 0 for those who continued.
Average School Certificate Grade at Age 15 (Ave_Grade)	The average value of all School Certificate (10th grade, age 15) Examination Marks over all subjects taken with weightings of 3 for an A, 2 for a B, 1 for a C and 0 for a fail (D).
Female	Binary=0 for a male; 1 for a female.
Maori	Binary=1 if Maori.
Pacific Islander (P_Island)	Binary=1 if a Pacific Islander.
Mother without Qualifications (M_NO_Q)	Binary=1 if child's mother does not have formal educational qualifications (10th grade School Certificate or higher).
Mother with Tertiary Qualifications (M_TERT_Q)	Binary=1 if a child's mother has a university or other tertiary qualification.
Father without Qualifications (F_NO_Q)	Binary=1 if a child's father does not have formal educational qualifications (10th grade or higher).
Father with Tertiary Qualifications (F_TERT_Q)	Binary=1 if child's father has a university or other tertiary qualification.
Number of Siblings (Num_Sib)	Number of siblings in the home at 15 years.

Table A1: continued

Average Income Decile (Inc_Decile)	Average income decile of the family when the child was between ages 11 and 14: 1 is consistently poor; 10 is consistently affluent.
Early Childhood Average Income Decile (Early_Inc_Decile)	Average income decile of the family when the child was between ages 1 and 5: 1 is consistently poor; 10 is consistently affluent.
Parents Own their Own Home (Own_Home)	Binary=1 if parents own their own home and the child is living at home at 15 years of age.
Rural Lifestyle (Rural)	Binary=1 if a child was not living in a main urban centre at 15 years of age.
Proportion of Family Income from Benefits (Ben_Prop)	The proportion (between 0 and 1) of the family's income derived from social welfare benefits.
Registered Unemployment (Local_Unemp)	Regional unemployment rate by gender in which each individual was living at 15 years of age. [20] There were 8 regions and the corresponding levels of unemployment ranged between 5.9 and 12.1 per cent.
Total Intelligence Quotient (IQ8)	The child's measured total IQ score testing cognitive performance at 8 years of age (revised Wechsler Intelligence Scale for Children).
Proportion of Students Continuing (Prop_Continue)	Proportion of an individual's 10th grade (Fifth Form) class within the data set continuing onto the 11th grade. The relevant individual is excluded from the calculation.
Class Size (Class_Size)	Average class size in secondary school
Affiliation with Deviant Peers (Peer_Dev)	Affiliation with deviant peers at age 15 based upon self-reported use of tobacco, alcohol, illicit drugs, other illegal behaviour, etc. by friends: 0-10, with 10 being the most deviant affiliations.

[20] Source of this information is the 1991 Census of Population and Dwellings: Regional Summary.

Table A2: Academic Performance

Single Equation Estimation (TOBIT):
(Equation (10): National Exam Grade at age 15)

Ave_Grade	**Coef.**	**t**	**P>|*t*|**
Dropout at age 16	-.4076	-3.39	0.001
Female	.2165	4.05	0.000
Maori	.1180	1.06	0.287
P_Island	-.1031	-0.59	0.553
IQ8	.0285	13.54	0.000
M_No_Q	-.1337	-2.09	0.037
M_Tert_Q	.1197	1.61	0.108
F_No_Q	-.0256	-0.41	0.681
F_Tert_Q	.1339	1.71	0.088
Inc_Decile (ages 11-14)	.0596	4.01	0.000
Early_Inc_Decile (ages 1-5)	.0413	2.92	0.004
Num_Sib	.0087	0.29	0.773
Own_Home	-.1480	-1.39	0.165
Rural	.0428	0.58	0.562
Ben_Prop	.2180	1.94	0.053
Prop_Continue	.5327	2.18	0.029
Peer_Dev	-.0587	-4.79	0.000
Constant	-2.7504	-8.75	0.000
Variance of error term	.6290	.0197 (std. error)	

Observation summary:
- 56 left-censored observations at Ave_Grade<=0
- 532 uncensored observations
- 10 right-censored observations at Ave_Grade>=3

Number of observations = 598 LR $chi^2(17)$ = 403.76

Prob > chi^2 = 0.0000

Log likelihood = -574.5929 Pseudo R^2 = 0.2600

Table A3: School Leaving

Single Equation Estimation (PROBIT):
(Equation (11): Dropout: 1=Left School at Age 16; 0=Enrolled in School at Age 16)

Dropout	Coefficient	z	P>$\|z\|$
Ave_Grade	-.7107	-4.14	0.000
Female	-.6974	-2.76	0.006
Maori	.3278	1.00	0.315
P_Island	-.2420	-0.47	0.639
M_No_Q	.2183	0.92	0.355
M_Tert_Q	-.1920	-0.51	0.608
F_No_Q	.1496	0.65	0.513
F_Tert_Q	.0595	0.16	0.871
Inc_Dec (ages 11-14)	.0106	0.17	0.867
Early_Inc_Dec (ages 1-5)	.0865	1.72	0.086
Num_Sib	.2430	2.18	0.030
Own_Home	-.4465	-1.48	0.139
Rural	.9643	2.56	0.010
Ben_Prop	.6586	1.87	0.062
Local_Unemp	.6383	1.77	0.077
Prop_Continue	-.6072	-0.70	0.485
Class_Size	.0356	1.27	0.203
Peer_Dev	.1002	2.58	0.010
Constant	-9.1960	-2.27	0.023

Observation summary:

Number of observations = 598	LR $chi^2(18)$ = 74.63
	Prob > chi^2 = 0.0000
Log likelihood = -101.4764	Pseudo R^2 = 0.2688