Understanding Barriers to Educational Attainment Among Canadian Youth

by

Kelly Foley

B.E.S., University of Waterloo, 1997

M.A., Carleton University, 2001

M.A., University of British Columbia, 2003

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES

(Economics)

The University Of British Columbia

(Vancouver)

December 2009

Abstract

In this dissertation, I explore barriers to educational attainment faced by Canadian youth from three different perspectives. The first paper, which is joint with David Green and Giovanni Gallipoli, uses an extended version of an unobserved factor model to investigate the factors that influence a youth's decision to drop out of high-school. Our results support three main conclusions. First, ability at age 15 plays an important role in dropping out. Second, parental valuation of education has a substantial impact on medium and low ability teenagers. Third, parental education has no direct effect on dropping out once we control for ability and parental valuation of education.

The second paper poses the question ”Can neighbourhoods change the decisions of youth on the margin of university participation?” I use the fraction of adults who have at least a Bachelors degree in a small geographic area surrounding a youth's home to proxy one source of potential peers and role models. Results suggest that at the mean neighbourhoods have a substantial impact on the likelihood of attending university.

I further explore this finding by estimating the marginal effect of neighbourhood characteristics at different points on the socio-economic distribution, and the distribution of reading test scores. One striking finding is that while university participation among youth from families with university degrees is unaffected by neighbourhoods, the marginal effect of neighbourhoods is largest for the high-skilled youth from lower socio-economic backgrounds.

The final paper evaluates the impact of a post-secondary preparatory course on high school achievement. BC AVID, a pilot project implemented with a random assignment design, encouraged students whose grade 8 grades were within the B

and C range to enrol in advanced classes and provided academic support through an elective course. I find that while the program increased the chances that students enrolled in an advanced math course, it also increased the likelihood that students failed the provincial examination. BC AVID did, however, help students pass their course work and as a result more students obtained credit for the advanced math course.

Table of Contents

List of Tables

List of Figures

Preface

An answer came to Katie. It was so simple that a flash of astonishment that felt like pain shot through her head. Education! That was it! It was education that made the difference! Education would pull them out of the grime and dirt...

Watching her children struggling up the stairs with their tree, listening to their voices, still so babylike, she got these ideas about education.

"Francie is smart," she thought. "She must go to high school and maybe beyond that. She's a learner and she'll be somebody someday. But when she gets education, she will grow away from me. Why, she's growing away from me now. She does not love me the way the boy loves me. I feel her turn away from me. She does not understand me. All she understands is that I don't understand her. Maybe when she gets education, she will be ashamed of me– the way I talk. But she will have too much character to show it. Instead she will try to make me different. She will come to see me and try to make me live in a better way and I will be mean to her because I'll know she's above me."

–*A Tree Grows in Brooklyn*, Betty Smith (1943)

Acknowledgments

I am particularly indebted to my supervisor, David Green, who is a generous teacher and insightful researcher. I am grateful for his ability to respond with equal patience to pressing concerns which range from ‘Can you help me pay for x?’ to ‘The variances are negative!’. I am also grateful for my very helpful committee members: Thomas Lemieux, and Siwan Anderson. Throughout the development of this project, I have also benefited from comments and suggestions from the entire labour group at UBC, participants of the micro-empirical seminar series and many other faculty and students in the department.

Chapter 3 has also benefited from the comments of participants at the Canadian Economics Association as well as at seminars presented at IZA, DIW-DC, Abt Associates, the University of Massachusetts-Boston, Memorial University, and the Copenhagen Business School. Participants at the Canadian Labour Market and Skills Researcher Network Annual Meeting and the Institute for Research on Poverty Summer Workshop provided excellent comments for Chapter 2

The statistical analysis in Chapters 2 and 3 was produced from Statistics Canada microdata. The interpretation and opinions expressed are my own and do not represent those of Statistics Canada. I would like to thank Lee Grenon and Cheryl Fu for all of their assistance in the BC Interuniversity Research Data Centers.

Access to data for Chapter 4 was granted by the Millennium Scholarship Foundation and the Social Research and Demonstration Corporation (SRDC). I would like to thank Reuben Ford and all of the staff at SRDC for their assistance.

Projects like this would not be possible without the participation of survey respondents and those willing to participate in pilot projects. I would like to thank all the participants in the BC AVID pilot project and the respondents to YITS.

Statement of Co-Authorship

I hereby declare that this thesis incorporates material that is the result of joint research undertaken in collaboration with David Green and Giovanni Gallipoli. The research presented in Chapter 2 is a result of that collaboration. I contributed in all aspects of the collaboration including identification of the research question, performing the research, data analysis and manuscript preparation.

Chapter 1

Introduction

So many aspects of well-being appear to be shaped by the educational attainment of individuals. It is not surprising that economists, along with other social scientists, are preoccupied by understanding the determinants of educational attainment. Why some individuals are highly educated while others fail to complete even secondary school would seem to be a straightforward question. Yet, after several decades of scrutiny by numerous scholars, many issues are unresolved and vigourous debates continue in the literature. In this dissertation, I contribute to that literature by examining specific barriers to educational attainment faced by some Canadian youth.

The theme of disadvantage associated with some family backgrounds connects the three main chapters of my dissertation. The question to what extent is access to education determined by circumstances of birth rather than actions and decisions under the control of individuals implicitly motivates the work in these chapters. Because of the association between education and individual benefits, the answer to this question can inform policy related to inequality and may influence how society chooses to redistribute the benefits of education. Moreover, if individuals access education because of family characteristics not associated with skill or ability, it may be possible for policy to achieve a more efficient allocation of the public resources that generate human capital.

The second chapter of my dissertation, co-authored with Giovanni Gallipoli and David Green, uses Canadian micro-level data from the Youth in Transition

Survey to examine the channels through which family socio-economic status and unobservable characteristics affect children's decisions to drop out of high school. First, we document the strength of observable socio-economic factors: our data suggest that teenage boys with two parents who are themselves high school dropouts have a 16% chance of dropping out, compared to a dropout rate of less than 1% for boys whose parents both have a university degree. We examine the channels through which this socio-economic gradient arises using an extended version of the type of factor model set out in Carneiro et al. (2003).

Our results support three main conclusions. First, cognitive ability at age 15 has a substantial impact on dropping out. Second, parental valuation of education has an impact of approximately the same size as cognitive ability effects for medium and low ability teenagers. A low ability teenager has a probability of dropping out of approximately .03 if his parents place a high value on education but .36 if their education valuation is low. Third, parental education has no direct effect on dropping out once we control for ability and parental valuation of education. Our results point to the importance of whatever determines ability at age 15 (including, potentially, early childhood interventions) and of parental valuation of education during the teenage years. We also make a small methodological contribution by extending the standard factor based estimator to allow a non-linear relationship between the factors and a covariate of interest. We show that allowing for non-linearities has a substantial impact on estimated effects.

In the third chapter, also using the Youth in Transition Survey (YITS), I estimate the relationship between neighbourhoods and university participation among Canadian youth. I use the fraction of adults who have a Bachelors degree in a small geographic area surrounding a youth's home as a proxy for the pool of potential peers and role models. Because the characteristics which determine where families live are generally correlated with university participation, I employ a conditional independence assumption to identify the neighbourhood effect. Specifically, I assume that, conditional on high school selection, family background, and a measure of literacy skills, selection into neighbourhoods is uncorrelated with the unobservable determinants of university participation. Results suggest that relative to the average youth living in a neighbourhood where no adults have a BA, a similar youth living in a neighbourhood where all adults have a BA is roughly 24 percentage points (or 60 percent) more likely to attend university.

Because the role that neighbourhoods play in the decision to attend university is probably small relative to other economic costs and benefits, one might expect that neighbourhoods will have a determining impact only for those on the margins of participation. I explore this idea by estimating the marginal effect of neighbourhoods at different points on the socio-economic distribution as measured by family income, parental education, parents' stated aspirations for their children's educational attainment, and the youths' scores on a literacy skills test. One of the most striking findings is that while university participation among youth from families with Bachelors degrees is unaffected by neighbourhoods, the marginal effect of neighbourhoods is largest for the most skilled youth from lower socio-economic backgrounds. This suggests that highly skilled youth from disadvantaged backgrounds are more likely to be affected by the negative influences in their neighbourhood relative to youth from families with high levels of education. Youth whose parents have a Bachelors degree attend university with above average probability, independent of both their literacy skills and any influences in their neighbourhood.

The fourth chapter evaluates the impact of a pilot project, called BC AVID, which was implemented in the province of British Columbia and was designed to help prepare middle-achieving high school students for post-secondary education. The program encouraged students to enroll in advanced classes and provided academic support through an elective course.

In the fourth chapter, I explore the impact of BC AVID on enrolling in an advanced grade 10 math course. The results presented in the chapter suggest that after two years students who were randomly selected to participate in AVID were more likely to enrol in the advanced course called Principles of Math. The program also increased the chances that students failed the provincial examination in Principles. Although they were less likely to be successful in the examination, BC AVID increased the chances that students passed their course work and as a result more students obtained credit for Principles. Non-experimental analysis suggests that relative to Principles students in the control group the marginal students who took Principles of Math because of AVID had similar scores on standardized tests taken in Grade 7. These students were more likely to be boys and less likely than their counterparts in the control group to aspire to obtain a university degree.

1.1 Bibliography

P. Carneiro, K. T. Hansen, and J. J. Heckman. Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review*, 44(2): 361–422, 2003.

Chapter 2

Ability, Parental Valuation of Education and the High School Dropout Decision[1]

2.1 Introduction

The strong correlation between family socio-economic status [2] and dropping out of high school is well known (e.g., Eckstein and Wolpin (1999), for the U.S. and Belley et al. (2008) for Canada). In the Canadian data we describe below, teenage boys with two parents who are themselves high school dropouts have a 16% chance of dropping out, compared to a dropout rate of less than 1% for boys whose parents both have a university degree. This cross-generational correlation is of interest for two reasons. First, to the extent it reflects higher ability individuals dropping out of high school, it may imply a social efficiency loss. Second, understanding this correlation is important when thinking about redistribution.[3] Our goal in this paper

[1]A version of this chapter has been submitted for publication. K. Foley, G. Gallipoli, and D.A. Green. Ability, Parental Valuation of Education and the High School Dropout Decision.

[2]Defined as encompassing parental education, parental income, and family structure.

[3]Under one theory of justice (Dworkin 1981a;b) luck is divided into 'brute luck' (outcomes for which an individual is not morally responsible, such as parental education) and 'option luck' (outcomes for which a person is responsible, such as effort exerted in pursuit of education). Differences arising out of brute luck call for redistribution, while differences due to option luck do not. Under-

is to provide a better understanding of the source of the socio-economic gradient in dropping out of high school.

There is a rich literature examining the high school dropout decision, particularly for the US. In papers dating back to the 1960s, researchers developed variants of what came to be called the 'Wisconsin Model' of educational and occupational attainment (e.g. Sewell et al. 1969, Alexander et al. 1975, Haveman et al. 1991). A key element of this model was its emphasis on the development of educational aspirations during adolescence and the importance of parents and peers in shaping those aspirations. Parental aspirations for their children were seen to be of particular importance (see for example Davies and Kandel (1981)). Recent work by Attanasio and Kaufmann (2009) also suggests that parental expectations and influences have a non-negligible effect on early education decisions. We provide further evidence that parental aspirations are strongly correlated with dropping out of high school. However, the interpretation of these results is complicated. Parents' answers to questions about the level of education they 'hope' their child will attain could reflect their own valuation of education in general, or an assessment of their child's own capabilities, or some combination of the two. If they reflect the former, this suggests policy responses targeting how parents or others influence students. If, instead, the answers reflect insider knowledge about a child's own abilities, then policies should focus on how to generate those abilities in the first place.

Within the more recent economics literature, Eckstein and Wolpin (1999) use a structural dynamic choice framework to examine dropping out in the US. They find that dropouts have lower ability and motivation as well as lower expectations about rewards from graduation. Todd and Wolpin (2006) investigate the form of the production function for cognitive skills using data from the NLSY79 Children Sample, focusing on implications for racial gaps in test scores. They find that mother's ability (as measured by the AFQT test) has a large impact on test score

standing how much of dropping out is due to elements beyond an individual's responsibility is then useful for thinking about the appropriate level of redistribution in society. Using James Heckman's words, we are interested in understanding the ramifications of the fact that there is 'no market for parents' (see Carneiro and Heckman 2003).

outcomes.[4]

Cunha and Heckman (2007; 2008), and Cunha et al. (2006) – hereafter, CH07, CH08, and CHS, respectively – investigate the production of both cognitive and non-cognitive abilities using dynamic factor models. As in Heckman et al. (2006), they find that both cognitive and non-cognitive skills are important determinants of dropping out of high school. They also include the type of home environment variables used in Todd and Wolpin (2006), but while Todd and Wolpin (2006) combine these variables in a fixed index, CH08 and CHS estimate weights to combine the variables into what they call a parental investment factor. They find that parental investments are important determinants of skill formation and, through them, of high school graduation, with investments having more impact on cognitive abilities at young ages and on non-cognitive abilities at older ages. We view the questions underlying these home environment or investment indexes as partly reflecting the parents' underlying valuation of education and the associated returns, defined in terms of both pecuniary and non-pecuniary benefits. Hence, this paper is an attempt to investigate the nature of family traits underlying the differences in parental investments which these earlier papers have shown to be an important determinant of high school graduation.

Our empirical approach builds on original results developed for factor based models by Carneiro et al. (2003) – hereafter, CHH – as well as CH07, CH08, and CHS. We carry out our investigation using the 'Youth in Transition Survey' (YITS), a rich longitudinal dataset in which a sample of over 20,000 Canadian teenagers are surveyed and given the PISA reading aptitude test at age 15 (in the year 2000) and then re-interviewed every two years thereafter. Our key dependent variable is whether the child is no longer in school and has not graduated at age 19. The YITS also includes a survey of the parents and a school administrator when the child is 15. It contains a long list of questions related to individual characteristics often seen as reflecting non-cognitive abilities as well as questions related to peers, the home environment and aspirations. As CHH argue, factor models provide an ideal vehicle for examining data of this type in which there are multiple noisy measures

[4] They also use home environment indexes which vary by age and which are built on answers to survey questions such as, whether the parents read to their child, how many books the child owns, whether the child has a musical instrument, and more.

of facets of interest. Following their approach, we set out a system containing a process determining dropout status augmented by a set of measurement equations related to the key underlying factors: cognitive ability, non-cognitive ability, and parental valuation of education.

Identification of the impacts of the factors in this class of models is obtained through covariance restrictions. CHH note that these restrictions can have an arbitrary quality and conclude that it is important to appeal to economics to justify any specific identification scheme. With this in mind, we begin our paper with a simple model of ability generation, parental valuation of education and the graduation decision. We use this model to guide our specification decisions. Our model has much in common with those in Todd and Wolpin (2006) and CH07. Unlike them, we do not have information on inputs and outcomes before the students are age 15 and we use the model to focus on the implications of different factors after that age.

Our empirical investigation proceeds sequentially through a series of estimators. We begin by documenting the unsurprising result that a student with less educated parents is more likely to drop out of high school - what we call the socio-economic gradient. We then present results from reduced form specifications which show that controlling for peer characteristics, proxies for non-cognitive ability, and school inputs does not alter the socio-economic gradient. However, including proxy measures of cognitive ability and parental valuation of education cuts the gradient in half. Given the potential problems with proxy estimators and difficulties with the interpretation of parental aspirations measures, we employ a system estimator in the spirit of CHH. It is at this point that the insights from the theory and the earlier literature become important. In particular, we argue that if the production functions generating the ability factors are assumed to be linear (which is the way they are implemented in CH08) then the factor estimator can be implemented with shapes for factor distributions which are common to everyone. This is the standard way to specify factor systems and when we estimate this model we find our results do not change substantially relative to the simple proxy estimator. However, we also follow CHS and allow the ability production functions to be non-linear by resorting to a more complex estimator in which the factor distributions can vary with parental education. When we implement this more flexible estimator, parental education ceases to have a statistically significant (or economically relevant) impact on dropping out. In other words, once we permit factor distributions to vary in shape as well as location, the differences in dropping out between teenagers from highly educated families and those from less educated families are completely accounted for by the fact that they are drawing from different ability and parental valuation distributions. The standard factor system estimator is a restricted version of our flexible estimator, and we test and reject those restrictions. We view this as evidence in favor of CHS's emphasis on non-linearities in the human capital production system. Our estimation approach allows us to account for these non-linearities in contexts where the researcher has less complete data than theirs - a not uncommon situation.

While parental education has limited effects on dropping out, the effects of cognitive ability at age 15 and parental valuation of education are substantial. The highest ability students are predicted never to drop out regardless of parental education or parental valuation of education. At the same time, parental valuation of education has a substantial impact on medium and low ability teenagers. A low ability teenager has a probability of dropping out of approximately .03 if his parents place a high value on education but .36 if their valuation is low. Non-cognitive ability has smaller impacts than either of these other factors.

The interpretation of the results from our most flexible estimator depends on the assumed underlying economic model. Recall that we observe cognitive and non-cognitive ability measures when the student is aged 15. If we assume a 'summary index' model in which all that is relevant about a child for assessing her probability of dropping out is her value of those ability factors, regardless of how they were generated, then our results imply that parental education has no effect on education decisions made in the teenage years (though it may have important impacts on generating cognitive and non-cognitive abilities at younger ages). Thus, there are no real gains to having parents who might be better able to understand your homework. On the other hand, having parents who care about education, perhaps investing time in children's education or just putting pressure on them to work hard as a result, has a substantial effect during the teenage years regardless of the parents' own education. The total effect of parental valuation of education would presumably be larger than this, since valuation is likely to play a role in determining ability at age 15. However, if we relaxed the summary index assumption (e.g., if there was an unaccounted for, unobserved factor partly determined by parental valuation of education) then these conclusions about the timing of factor impacts could be overstated. Given the wealth of measures we are able to control for, we favor the single index interpretation. But, in either case, our results point to the importance of whatever determines ability by age 15 (including early childhood interventions) and of parental valuation of education. We view these results as potentially hopeful since they suggest we might affect dropout rates in ways other than the slow, cross-generational process of raising parental education. Instead, policy could focus on replicating whatever high valuation parents do for their children. This points to the importance of studies looking at what high and low educated parents do differently, for example Carneiro et al. (2007).[5]

The paper is organized as follows. In Section 2.2 we present a simple life-cycle model describing the decision to drop out from high school. In Section 2.3 we map the model to data, setting out an estimable counterpart. In Section 2.4 we describe the data, and Section 2.5 contains results from the various estimators described earlier. In Section 2.6 we summarize and conclude.

2.2 A Simple Decision Problem

In this section, we set out a model of the decision to drop out of high school in an inter-temporal optimizing framework. Teenagers are assumed to make the dropout decision rationally based on expected returns given their levels of ability and their information on the returns to education. We recognize that modeling teenagers as rational, forward-looking agents may stretch credulity to some extent so we also modify the model to allow parents to enforce a minimum effort level. Our goal is to use the model to illustrate key issues in dropout determination and to obtain guidance for setting up and interpreting our empirical specifications.[6]

In setting out the model, we divide individual lives into three periods, numbered from zero to two. The middle, or teenage, period (period 1), corresponds to

[5]In this paper, we do not investigate the channels through which parental valuation operates.

[6]Given the relatively low monetary cost of getting access to high schools in Canada, we do not explicitly model credit constraints in the model. In the empirical analysis, however, we control for both short- and long-term constraints.

the time after the legal school leaving age (16 in most Canadian provinces in our sample period) and before the typical graduation age (18). The dropout decision is made in this period and we model it as conditional on the ability the teenager has accumulated in period 0 (i.e., up to age 15) and on expected returns to high school graduation in the future (period 2). We assume that the student does not make optimizing decisions in period 0 but we begin with a description of that period because assumptions related to the generation of ability in that period are relevant for the interpretation of our estimates.

2.2.1 Period 0: The 'Shaping' of Teenagers

We assume that a child is endowed with an ability vector, θ_0, at birth. The vector has two elements, corresponding to cognitive and non-cognitive ability and is determined by,

$$\theta_0 = f_1(\theta_F, \iota) \tag{2.1}$$

where $f_1(.,.)$ is a (possibly non-linear) function, θ_F is a (2×1) vector of hereditary cognitive and non-cognitive abilities characterizing a family and ι is a vector of individual-specific traits which are randomly assigned.

Ideally, we would like to separately account for youth's ability and observable parental characteristics, such as education. However, this is complicated by:

1. The fact that parental education is likely a function of θ_F. Indeed, we will assume that parental education (PE) is determined as,

   $$PE = f_2(\theta_F, \nu_p, \eta) \tag{2.2}$$

   where ν_p corresponds to parental valuations of the return to education and η summarizes all factors contributing to PE which are orthogonal to θ_F and ν_p.

2. The ability we observe in the data - ability at the start of the teenage period, denoted as θ_1 - will itself be a function of parental inputs. In particular, we assume it is a function of θ_0, of parental education (either because it reflects family income effects or because hours of parental time from educated parents are more effective in generating children's ability) and of parental

valuation of education (because it helps determine how much effort parents invest in improving their child's ability). That is,

$$\theta_1 = f_3(\theta_0, PE, \nu_p) \tag{2.3}$$

where, following CH07, it is possible that cognitive ability at age 15 is a function of both endowed cognitive and non-cognitive abilities, and the same is true for age 15 non-cognitive ability.

We treat the dropout decision in the teenage years as conditional only on the actual value of θ_1, not on the specific combination of innate ability and family inputs that generated that value. If we assume, in addition, that there are no further factors that are both relevant for education decisions and a function of ν_p and PE, then the vector θ_1 is a sufficient statistic for all the education related decisions that were made before the teenage years. Under these assumptions, the estimated effects of parental education and ν_p are interpreted as effects not already reflected in θ_1, that is, as new effects on the dropout decision after age 15.

2.2.2 Periods 1 and 2: The Dropout Decision and After

In period 1, a teenager has two options: study toward a high school degree or work at the market wage for dropouts (denoted as wage w_{LHS}). [7] In period 2 (representing the remainder of life), dropouts earn w_{LHS}, so that their discounted value of lifetime utility is $U(w_{LHS}) + \beta U(w_{LHS})$. The period 2 earnings for graduates are higher and we will assume they are determined by,

$$w_2^p = \alpha_0^p + \alpha_1 \theta_1 + \alpha_2 grd \tag{2.4}$$

where grd is a measure of academic performance, α_0^p and α_2 are scalars and α_1 is a vector.

The superscript 'p' in w_2^p indicates that the above equation represents a prediction conditional on the information available to the teenager and his or her family. We allow the information about market returns to differ across families and

[7] To simplify discussion we assume that w_{LHS} is the minimum wage and, therefore, is not a function of a person's abilities.

youths[8], by specifying,

$$\alpha_0^p = \alpha_{01} + \alpha_{02} PE + \alpha_{03} \upsilon_p \tag{2.5}$$

where v_p corresponds to parental predictions about returns to education.[9] Thus, children's notions of the returns to education increase with their parents perceptions of the same. Parental education is included on the assumption that more educated parents may have better information on the returns to education (Junor and Usher 2003). Note that while this specification incorporates predictions about future returns, the model still doesn't have any important uncertainty - each family acts as if it knows the returns to education; it's just that what they claim to know differs across families.

The academic performance measure in equation (2.4) is determined by,

$$grd = \psi_0 + \psi_1 e + \psi_2 PE + \psi_3 x_s + \psi_4 \theta_1 + \psi_5 \upsilon_p \tag{2.6}$$

where, the ψ's are parameters or vectors of parameters as required. Thus, academic performance is potentially determined by school inputs, x_s, the child's abilities, his or her effort level, e, and by parental inputs determined by PE and υ_p. The combination of (2.4) and (2.6) implies that the return to effort in school comes in the form of higher earnings in period 2. We do not explicitly model the related choices and outcomes in period 2 but this could arise, for example, if higher high school grades raise the probability of going to university, with its attendant higher earnings. Equation (2.4) can be seen as a linearization of such processes.[10]

2.2.3 Utility

We assume linear utility of consumption $U(c) = c$, in order to focus on expected pay-offs. Effectively, an agent chooses a consumption level c_t in each period $t \in \{1, 2\}$ by choosing whether or not to stay in school in period 1. Labour is

[8] Notice that w_2^p assumes values on some interval $[\underline{w}, \bar{w}]$ and that predicted education outcomes are approximated, at age 15, by equation (2.4).

[9] Such returns can be also non-pecuniary, and pertain to the perception that education is a 'good' in itself.

[10] It is possible to show that this linearization is consistent with a richer, less restrictive model of earnings generation. Details are available from the authors.

inelastically supplied in each period. The labour endowment in the second period, n, reflects the expected length of working life after age 18. Labour income is consumed in full during each period and agents have no means of transferring wealth between the two periods, with the noticeable exception of completing education. We assume that student consumption in period 1 is based on transfers from their parents determined by a combination of parental permanent income, PI, and current family income FI. This allows transitory income shocks to have an impact on education decisions as one might expect in the presence of credit constraints (Coelli 2009).[11] Students optimally choose 'schooling effort' e, which has a direct impact on their future earnings. Effort affects utility negatively and, for convenience, we assume that it enters utility additively through a general function $g(e) = -\left(\gamma e + \frac{1}{2}e^2\right)$, with $-\gamma > 0$ being a minimum level of effort. We assume that minimum effort level is a positive function of υ_p, implying that even myopic students may supply enough effort to graduate if their parents value education highly, perhaps because they gain utility through kudos from their parents. [12]

2.2.4 The Value of Working vs the Value of Studying

Lifetime utility for an agent who does not drop out can be written as

$$V_S = \max_e g(e) + f_S(PI, FI) + \beta n w_2^p \tag{2.7}$$

where $f_S(PI, FI)$ is a function of parental permanent income and current income denoting consumption by a youth while in school, β is the discount factor and n denotes expected working life. Lifetime utility for an agent who drops out of high school in period 1 is simply

$$V_W = \max_e g(e) + f_W(PI, FI) + (1 + \beta n) w_{LHS} \tag{2.8}$$

where $f_W(PI, FI)$ summarizes consumption transfers to a dropout youth in period 1.

[11] Students cannot increase their period 1 consumption by working.

[12] We can also allow γ, the utility cost of extra effort, to be lower if the student's peer group is more academically oriented, which can be captured through a vector of peer characteristics, z.

When deciding whether to drop out ($d = 1$) or not ($d = 0$), an agent compares V_S to V_W. In the absence of randomness, the problem can be written as

$$\max_{\{d\}} (1-d)V_S\left(e^*_{1-d}\right) + dV_W\left(e^*_d\right) \quad \text{with } d \in \{0,1\}$$

that is, conditional on all states in the problem, an agent optimally chooses the value of studying or, alternatively, working and the implied effort level. Given the objective functions and constraints, we can easily show that the optimal effort of a student is

$$e^*_{1-d} = \beta n \psi_1 \alpha_2 - \gamma \tag{2.9}$$

Optimal effort is a function of patience and expected working life of an agent, as well as of the relative importance of effort in determining schooling outcomes[13]. It will also be affected by parental valuations of the returns to education and peer characteristics through their impacts on γ.

2.2.5 Empirical Specification for the Dropout Decision

The decision of whether to drop out from high school is determined by the difference between the lifetime utilities associated with dropping out and graduating, each evaluated at the corresponding optimal effort level. Assuming that $f_S(\cdot)$ and $f_W(\cdot)$ are linear and that parental permanent income, PI, can itself be approximated as a linear function of parental education, this difference is given by,

$$I_D = \gamma_0 + \gamma_1 FI + \gamma_2 PE + \gamma_3 w_{LHS} + \gamma_4 x_s + \gamma_5 z + \lambda_{d\theta 1}\theta_{11} + \lambda_{d\theta 2}\theta_{12} + \lambda_{d\upsilon}\upsilon_p + u_0 \tag{2.10}$$

where, θ_{11} and θ_{12} correspond to cognitive and non-cognitive ability, respectively, and u_0 is an error term that incorporates an idiosyncratic component of current utility as well as any added randomness associated with the grade function and second period earnings for graduates. This index function completely determines dropping out, with d=1 iff $I_D > 0$, and it is the basis of our estimation. Notice that because we have substituted in for optimal effort, variables such as hours of

[13] Notice that, conditional on choosing to drop out, the effort e^* is set at the minimum level $e^* = -\gamma$.

studying do not belong in the index.

Our main interest is in estimating γ_2, showing the impact of parental education on dropping out. Based on the model, that coefficient reflects an effect of parental education on grades attainment and on the student's evaluation of returns to education as well as proxying for family permanent income effects. Assuming a summary index type model (i.e., one in which θ_1 fully captures all factors relevant for the dropout decision from before age 15), this coefficient captures those effects going forward from age 15. Thus, if parental education only affects ability generation for young children then it would help determine θ_1 but $\gamma_2 = 0$. Estimation of (2.10) is complicated by the fact that we do not directly observe θ_{11}, θ_{12} or υ_p. In the next section, we present a series of empirical approaches to address this problem, using the model to help interpret what we obtain from each approach.

2.3 Empirical Strategies

2.3.1 Reduced Form

The simplest approach to estimating (2.10) is to ignore θ_{11}, θ_{12} and υ_p and implement (2.10) as a simple Probit without including measures of these factors. We do this by including a full set of province dummies to capture variation in minimum wages (w_{LHS}) and other provincial level policies. We also broaden our definition of socio-economic background by allowing for differences in dropping out by family structure, and examine impacts of measures of school inputs, peer characteristics, and local youth unemployment rates. As discussed earlier, the estimated *PE* coefficient will reflect the effects captured in γ_2 plus unobserved ability and parental education valuation effects.

2.3.2 Proxy Estimator

We can attempt to isolate γ_2 by introducing proxies for θ_{11}, θ_{12} and υ_p into our estimation. Our data includes results for students taking the PISA tests at age 15 (which we describe in more detail in the data section). We assume the test score is generated according to,

$$PISA = \delta_{10} + \lambda_{T\theta 1}\theta_{11} + u_1 \tag{2.11}$$

where, *PISA* is the PISA test score. In this equation, and the other measurement equations that follow, the δs and λs are either parameters or vectors of parameters, as required, and the *us* are error terms which are assumed to be independent of covariates, the factors and the error terms in all other equations. Equation (2.11) says that the test score is a true reflection of cognitive ability at age 15, observed with error. Because PISA is a one-time test, we assume the student's test score is not directly determined by effort, parental inputs, peer effects et cetera, except to the extent that they have shaped the student's ability on the test day, θ_1. This identifying assumption is important for the more complicated estimators we use later.[14]

As a proxy for υ_p, we will use a variable built from parents' responses to a question about the level of education they hope their child will achieve. We will call that variable $parpref$ and assume it is determined according to,

$$parpref = \delta_{20} + \delta_{21}PE + \lambda_{p\theta 1}\theta_{11} + \lambda_{p\theta 2}\theta_{12} + \lambda_{p\upsilon}\upsilon_p + u_2 \tag{2.12}$$

Choosing a proxy for the non-cognitive element in the ability vector is complicated by the fact that non-cognitive abilities are heterogeneous and difficult to reduce to one factor. Borghans et al. (2008) argue for classifying these abilities into the Big Five factor scheme favored by some psychologists. However, they also present evidence that among the Big Five factors, Conscientiousness is strongly related to education outcomes while several of the others are not. Rather than try to extract a factor from a disparate set of questions, we restrict ourselves to questions related to Conscientiousness. Conscientiousness relates to being achievement oriented, self-disciplined and confident. As a primary proxy for this, we use a question asking the student how often the statement, "I do as little work as possible. I just want to get by," is true for him or her.[15] We code a variable equaling 1 if they

[14]This variable and all the other measurement variables related to the factors occur as categorical variables in our data so these equations should be interpreted as index functions underlying the actual realizations of the measurement variables.

[15]The YITS dataset includes also a self-efficacy index which is potentially useful since self-efficacy is related to Conscientiousness. However, the questions underlying this index relate to

answer 'Never' and assume this is determined by an underlying index function,

$$getby = \delta_{30} + \delta_{31} PE + \lambda_{c\theta 2} \theta_{12} + \lambda_{c\upsilon} \upsilon_p + u_3 \tag{2.13}$$

Whether a child provides only the minimum effort depends on their level of conscientiousness (θ_{12}) but also on parental valuation of education since parents who value education highly may pressure children to do more than the bare minimum.[16]

Using proxies may reduce our identification problems but will likely not eliminate them for two reasons. The first is the well-known problems with endogeneity that arise when using proxies of this type. Thus, if we solve (2.11) for θ_{11} and substitute into (2.10), we obtain an estimating equation with $PISA$ on the right hand side but also with the disturbance that helps determine it, u_1, in the error. Thus, estimates will be inconsistent. In particular, the coefficient on PE will still reflect ability effects to the extent that the part of ability not fully captured in the test score is correlated with PE. Given our assumption that PE helps generate θ_{11}, it seems likely that such a correlation exists.

The second issue is with interpretation. We are interested not only in the coefficient on PE but also in the effects of θ_{11}, θ_{12} and υ_p themselves. As indicated in (2.12), it seems plausible that parental responses to a question about the level of education they hope their child will achieve will reflect how much they value education but may also reflect their evaluation of the child's ability. That is, a parent who knows his or her child's true values of θ_{11} and θ_{12} are low may set lower expectations for that child. To the extent that $PISA$ mismeasures ability, the coefficient on $parpref$ may partly reflect the insider knowledge of the parent about the child's true abilities rather than just υ_p.

whether the student thinks he or she can do well on tasks at school, which appears to be as much a self-assessment of cognitive abilities as a measure of self-efficacy so we use it only in our reduced form.

[16]Note that, following CH08 and CHS, we assume that the current value of this measurement variable reflects only non-cognitive ability, θ_{12}, though cognitive abilities may have been an input into the production of θ_{12} itself in the past.

2.3.3 Unobserved Factor System Estimators

Basic Estimator

We can potentially solve the problems with the simple proxy estimator by using further information related to ability and parental valuations. Like many panel datasets, the YITS includes a large set of variables, with the number expanded by the fact that parents and children are asked separate sets of questions. CHH propose using extensive sets of variables such as these to construct a system of measurement equations in the spirit of factor analysis to identify and control for the effects of latent factors. As they discuss in detail, identification of the effects of these factors and of parameters related to their distribution requires that we have at least two such measurement equations related to each factor, along with the main estimating equation, (2.10). The test score equation (2.11), provides one such measurement equation related to cognitive ability. Another natural candidate for this is the equation determining grade 10 grades. This can be obtained by substituting the expression for optimal effort into equation (2.6), yielding,

$$grd = \delta_{40} + \delta_{41} PE + \delta_{42} x_s + \delta_{43} z + \lambda_{g\theta 1}\theta_{11} + \lambda_{g\theta 2}\theta_{12} + \lambda_{g\upsilon}\upsilon_p + u_4 \qquad (2.14)$$

To capture parental aspirations for education, we use the *parpref* equation plus an equation corresponding to parents' answers to a question about whether they have saved for their child's future education. We use this as a dummy variable the value of which is determined by an underlying index function,

$$saved = \delta_{50} + \delta_{51} PE + \delta_{52} FI + \lambda_{s\theta 1}\theta_{11} + \lambda_{s\theta 2}\theta_{12} + \lambda_{s\upsilon}\upsilon_p + u_5 \qquad (2.15)$$

Thus, holding family income constant, parents who value education more highly are more likely to save for their children's education. As with the *parpref* variable, savings behavior may partly reflect parents' information on their child's ability.

We measure non-cognitive ability using the *getby* variable plus a variable based on a question asking the student whether she completes her assignments. This is related to the organization and goal-oriented dimensions of Conscientiousness. We specify the index function determining this variable as,

$$hmwork = \delta_{60} + \delta_{61}PE + \lambda_{h\theta 2}\theta_{12} + \lambda_{h\upsilon}\upsilon_p + u_6 \tag{2.16}$$

where we have again assumed parents have an effect on achieving education related outcomes such as handing in homework.

Together, equations (2.10) through (2.16) constitute a system with the dropout process specified jointly with measurement equations that will help identify the ability and parental education value factors in the dropout process. CHH discuss the conditions under which one can obtain identification for all the factor loadings on θ_{11}, θ_{12} and υ_p in these equations as well as the parameters which define the distributions for θ_{11}, θ_{12} and υ_p. In particular, in our system we can obtain identification if one of the measurement equations includes only one of the factors. This is satisfied by the *PISA* equation, which we argued plausibly includes only the θ_{11} factor. We also need to normalize one of the loadings for each factor to one. We set $\lambda_{T\theta 1}$ (the loading on θ_{11} in the *PISA* equation), $\lambda_{s\upsilon}$ (the loading on υ_p in the *saved* equation) and $\lambda_{c\theta 2}$ (the loading on θ_{12} in the *hmwork* equation) to one. With these restrictions and assuming the factors are mean zero and orthogonal to one another, we have 17 parameters related to the factor distributions to identify (counting the factor loadings that have not been normalized to one plus the variances of the factors). We have 19 unique covariances which are allowed to be non-zero in the structure among the errors of the dropout equation and the 6 measurement equations.[17] Thus, the order condition for identification is met. The rank condition corresponds to whether the specific pattern of entry of factors in the various equations allows us to recover all 17 parameters. This is indeed the case.[18]

Examining expressions determining the various parameters as functions of observable covariances provides some insight into the source of identification. For example, suppose that we set $\lambda_{g\theta 1}$ (the factor loading on θ_{11} in the *grd* equation) to one, placing no restrictions on the factor loadings for θ_{11} in the other equations. It is simple to show that the expression for $\lambda_{d\theta 1}$ (the factor loading on θ_{11} in the

[17] As CHH discuss, identification of the coefficients on the observable variables is given and so we can discuss identification of the factor loadings and variances in terms of the dependent variables net of the effects of the right hand side variables - i.e., the broadly defined errors in all the equations.

[18] A document demonstrating a solution is available upon request.

dropout process) is given by,

$$\lambda_{d\theta 1} = \frac{cov(I*_G, T*)}{cov(grd*, T*)} \tag{2.17}$$

where the '*' indicates that we are discussing covariances after netting out the effects of observable covariates, and, following CHH, we discuss the index corresponding to the dropout decision as if it were an observable, continuous variable. Equation (2.17) is immediately recognizable as the instrumental variable estimator one would obtain if *grd* were included as the right hand side variable in the dropout equation and *PISA* were used to instrument for it. Thus, this estimator is in the same spirit as Chamberlain's (1977) estimator for the impact of schooling on earnings in which the outcomes of two other family members are used to address an unobservable family ability factor. The variation being used in this estimator is essentially the parts of *grd* and *PISA* that they have in common.

The result in (2.17) suggests that we can obtain consistent estimates with a simple instrumental variable (or control function) estimator. However, we would like to allow for a flexible form for the distribution of the error in the dropout process and to use an estimator that permits an extension that we detail in the next subsection. For both reasons, we turn to an estimator in the spirit of Heckman and Singer (1984). In particular, we represent θ_{11}, θ_{12} and υ_p as having discrete distributions that are independent of one another. Further, we assume that the errors in the system, $u_0, ..., u_6$ are all normally distributed and independent of one another and of the factors. Thus, conditional on specific values for the factors, an individual's contribution to the likelihood function is just the product of normal CDF evaluations (since all the dependent variables are actually discrete). This product is calculated for each possible combination of values for the factors, then these factor conditional products are each multiplied by the associated probability of observing that set of factor values and then summed.[19] The factors provide a flexible way to link the various equations, representing the joint distribution as a flexible mixture of normals. Maximizing the likelihood function provides estimates of the γ and δ vectors as well as the factor loadings ($\lambda's$), and the locations of the points of support and the associated probabilities for the factor distributions. Most importantly,

[19]We discuss the likelihood function in Appendix B.

assuming that equations (2.1) through (2.3) are linear (and, therefore, the dropout process and all the measurement index functions are linear in the factors), this system provides consistent estimates of γ_2 and the other parameters of interest. It is worth noticing that it does so in the context of a system in which we explicitly allow for the possibility that observable measures of parental educational aspirations partly reflect parental insider knowledge about the abilities of their children.

Extended Estimator

The assumption that all relevant equations are linear in the underlying factors is potentially strong. In particular, the results in Gallipoli et al. (2009) suggest that the relationship between parental and child ability is non-linear and CHS argue for non-linear versions of the skill production function, (2.3). To understand the implication of these non-linearities for our estimation, note that we are interested in characterizing the conditional distribution,

$$p(Y,\theta|PE;\Gamma) = \int p(Y|\theta,PE;\Gamma)p(\theta|PE;\Gamma)d\theta \tag{2.18}$$

where Y is the matrix of outcomes (both dropping out and the measurement values), θ is a vector containing all the factors, and Γ is the matrix of all parameters in the system, including those defining the factor distributions. Our problem relates to the impact of parental education (PE) so we have written the probability as conditional upon PE but suppressed other covariates. If the factor generation equations (2.1)–(2.3) are linear then the shape of the factor distribution is the same for all values of PE and we can write the likelihood with the factor distribution estimated unconditional on PE, i.e., using $p(\theta;\Gamma)$. This is the form of the standard version of this estimator. However, if, for example, θ_1 is a nonlinear function of PE then the correct specification of the likelihood function is given in (2.18) with the factor distribution being conditional upon PE. If this is the case, but we implement the more standard version of the estimator, then the effect of PE in determining the shape of the θ distribution could be reflected in the coefficient on PE in the dropout and measurement equations.[20]

[20] To verify this, we constructed a Monte Carlo exercise in which we generated values for a single factor and for PE based on equations (2.1)–(2.3). We then generated values for dropping out based on an index expressed as a linear function of PE and the factor and also generated values for two

To address this issue, we specify and implement an 'extended' factor estimator in which the points of support for the factor distributions are the same for every observation but the probabilities associated with those points are allowed to differ by parental education. This introduces an additional channel for parental education to affect dropping out: through this channel, two students with the same abilities and parental valuation will have the same probability of dropping out even if their parents have very different education levels. However, their 'ex-ante' probabilities of having those factor values could be very different. Thus, this estimator allows for the possibility that estimated PE effects in the linear system estimation are partly disguised ability and parental valuation effects that are loaded onto the PE coefficients because PE and the factor distributions have a more complex relationship than is allowed in the basic estimator. Identification of the different probability weights stems from the extent to which the distributions of our factor related measures (grades and PISA score for the ability factor) do something other than simply shift proportionally when parental education changes.[21]

2.3.4 Interpreting Parental Valuation

One of our main interests is in the potential impact of parental valuation of education. As we discussed earlier, trying to uncover this impact is complicated by the fact that questions asked of the parents about preferences for their children's education or about their investments in that education may reflect 'insider' knowledge about their children's abilities. In our model and estimation, we explicitly allow our parental valuation measures ($parpref$ and *saved*) to be related to child cogni-

proxies for the factor, measured with error. When we specify (2.1)–(2.3) as linear equations we find: 1) a proxy estimator using one of the proxies resulted in estimates of the coefficient on PE in the dropout equation that were biased upward from their true value; 2) both an IV-type estimator in which we used the second proxy as an instrument for the first and the standard system estimator generated unbiased estimates of the coefficient on PE. However, when we allowed (2.1)–(2.3) to be nonlinear, the proxy, IV, and standard system estimator all produced upwardly biased estimates of the PE coefficient.

[21] We have also extended the basic estimator by allowing the two factors to be correlated rather than orthogonal. This is in the spirit of 'oblique' factor models used in other parts of the social sciences and is something allowed in CH08 and CHS. The conclusions from the estimates with this specification are not substantially different from those in which the factors are assumed to be orthogonal. Because the orthogonality allows easier interpretation, we present our results using that specification and omit the correlated factor estimates to save on space.

tive and non-cognitive abilities. Thus, to the extent our parental valuation factor is picking up an unobserved ability, it must be an ability that is orthogonal to cognitive ability (as reflected in *Pisa* and *grades*) and non-cognitive ability (as reflected in *getby* and *hmwork*). Previous evidence suggests that cognitive abilities can be well (even if not perfectly) captured by one factor (Borghans et al. 2008). Based on this (and given how important parental valuation turns out to be in the dropout process), it seems unlikely to us that v_p is picking up further cognitive abilities. There is more variety in non-cognitive abilities. In the proxy specification in the next section, we include a number of individual variables and indexes related to non-cognitive traits such as self-esteem and self-efficacy. Their inclusion has little effect on the $parpref$ gradient in dropping out relative to just controlling for our main cognitive and non-cognitive proxies. Thus, if v_p is capturing some other unobserved ability, it must be an ability other than what is reflected in the extensive set of variables in the YITS. Based on this, we maintain an interpretation of the v_p factor as capturing parental valuation of education.

2.4 Data

We use data from the *Youth in Transition* (YITS) survey. The YITS is a longitudinal survey that tracks the experiences of two cohorts of Canadian youth. It provides a rich panel of information on the participants' demographic background, their participation in education and work, as well as their beliefs, attitudes and behaviours. The youngest cohort was 15 years-old when the first cycle of data was collected in 2000. Because schooling is legally required for age-15 children we use data from this cohort. The first cycle of data therefore provides a way to characterize a 'baseline'. In the YITS, participants are surveyed every two years. We also use data from the second and third cycles when the youth were 17 and 19 years-old.

The original sample of 29,687 students was drawn from a two-stage sampling frame. Schools were sampled first from a list compiled by provincial Ministries and Departments.[22] In the second stage, students were sampled within the 1,187 schools.[23] The sample size within each school was chosen to facilitate school-

[22]Schools were sampled from within the strata of province, age-15 enrollment size, linguistic group, public vs. private funding sources, and urban vs. rural settings.

[23]Schools were excluded from the sample if fewer than 3 students were present or likely to re-

level analysis. Because some provinces and linguistic groups were over-sampled, the within-school sampling rate ranged from less than 10 percent to a census of the 15 year-olds. In all of the results we report, we use weights provided by Statistics Canada that account for over-sampling, non-response to the parental survey, and longitudinal attrition. Approximately, 13 percent of the sample is lost due to non-response to the parental survey. The overall response rate to the third cycle was 66 per cent. Some cases were also lost due to missing data or invalid responses to questions. The final sample is 7,755 boys and 8,376 girls.

The YITS is useful, in part, because it includes a parents' survey completed by the parent or guardian who identified him or herself as 'most knowledgeable' about the child. The responding parent provided data about their and their partner's education, work, and income. Parents also answered questions about their attitudes toward and aspirations for their children.

At the time of the survey, the children also completed a reading test which was administered through the Programme for International Student Assessment (PISA). PISA was an effort, co-ordinated by the OECD, to generate internationally coherent measures of cognitive skills. We use data from the PISA reading cohort.[24]

We identify individuals as high school dropouts if, according to their self-report, they had not completed the requirements for a high school diploma and were not in school at the time of the Cycle 3 survey. The third wave of the YITS data was conducted between February and June 2004, when respondents were all age 19. In most provinces, this corresponds to the spring of the year following their normal graduation year with two notable exceptions. The first is Ontario, where there was an option for students to stay in school for an extra year (grade 13), allowing them more time to complete courses that would prepare them for university. For students born in the first half of the year, some could still be in school without ever having an interruption in their schooling at the time of the cycle 3 interview. These would have to be students who have not already dropped out or chosen to graduate after 12 years, are in the last term of high school, and are likely interested

spond to the survey. Schools for children with severe learning disabilities, schools for blind and deaf students and schools on First Nations reserves were also excluded.

[24]While the YITS project also includes science and math skills tests, we use reading scores because the sample is twice as large. All of the students wrote the reading test, and half wrote either the math or science test.

in attending university. Thus, it seems unlikely that many of them will ultimately be dropouts. The other exception is Quebec where high school ends at grade 11 and students interested in university then go to two-year preparatory schools called CEGEPS. We re-estimated our reduced form and proxy specifications after dropping all observations from Ontario or Quebec [25] and obtained very similar results to those presented here. Thus, we do not believe these two anomalies are driving any of our results. [26]

The unconditional dropout rates at age 19 in this data using our dropout definition are .055 for boys and .036 for girls. These compare with numbers from the OECD showing that 11 per cent of 20 to 24 year old Canadians (both genders combined) have not completed high school and are not currently in school (de Broucker P. 2005). Our rates could be lower than the OECD numbers, in part, because some students who have not yet graduated at age 19 but are still in school will ultimately drop out, causing the dropout rate to be higher in the 20 to 24 year old age window in the OECD data. Belley et al. (2008) also use YITS data and find that at the fourth (age 21) wave, the dropout rate for both genders combined is .07. We focus on dropping out at age 19 because we believe it provides a clearer picture of the role of family supports on the dropout decision and because it reduces the amount of sample attrition we face. Lower dropout rates in the YITS could also relate to dropouts being more likely to attrite from the sample. Sample weights used in all of our calculations are supposed to account for this but may not do so completely.[27] Finally, to place Canada's experience in context, Belley et al. (2008)

[25] It was possible to do this without the resulting sample being too small to use because the YITS strongly under-sampled Ontario and Quebec.

[26] Our dropout definition differs from the one used by some other authors (e.g. Eckstein and Wolpin 1999), who include current students who have not graduated as dropouts. We view counting these on-going students as dropouts as a potential mis-labeling that could cause us to miss relationships such as parents pushing their children to complete their schooling in "whatever time it takes." We re-estimated our model using Eckstein and Wolpin's definition and found similar results to those presented here with the main exception that the importance of parental valuation of education is somewhat reduced, though still economically substantial and statistically significant.

[27] Student attrition between cycles 1 and 2 and between cycles 2 and 3 implies that we have approximately 70% of the original sample with usable information by the third cycle. This is not an inordinately high attrition rate by the standards of most panel data. The weights provided to address this issue are essentially the outcome of an estimated attrition process. Thus, the variables used in constructing the weights can potentially play the role of exclusion restrictions. If there are variables used in constructing the weights that are not used in the final estimation then those variables effec-

use the NLSY to show that with a comparable definition of dropping out at age 21, the US dropout rate is .17, that is .1 higher than for Canada.

We describe our other variables as they arise in our estimation. A table of sample means is provided in Appendix A.

2.5 Results

2.5.1 Reduced Form

We begin with results from a reduced form Probit specification. This is useful, in part, because it allows us to establish the size of the socio-economic gradient we are trying to explain. As discussed in section 3, though, the coefficients on socio-economic variables such as parental education in this exercise will reflect both the direct effects of these variables and the effects of omitted ability and parental aspirations factors. In all the specifications that follow we include (but do not report) province indicators. Standard errors are calculated incorporating clustering at the high school level in all specifications because of the nature of the sampling scheme.

We first introduce a series of variables capturing the socio-economic status of the child's family. Key among these variables are parental education and income. Income is defined as total before-tax family income including transfers, expressed in thousands of dollars, and put into adult-equivalent form by dividing by the square root of the number of people in the family. Parental education is captured with a set of six categorical variables corresponding to the highest level of education achieved by both parents: 1) both parents are high school dropouts; 2) one parent is a high school dropout and the other is a dropout; 3) both parents are high school gradu-

tively become instruments for addressing selection. The information we have obtained from Statistics Canada suggests that the variables used in constructing those weights are all variables that are either included in our final specifications or are strongly related to included variables. One exception to this is a variable based on a question to the parents about whether they were willing to have their data shared with another government department (HRDC). We tried, unsuccessfully, to get access to this variable to allow us to model attrition explicitly ourselves. Instead, we tried implementing an estimator including an explicit attrition process and using a variable equalling the proportion of times a respondent did not answer a question asked of everyone as an exclusion restriction. However, this variable did not perform well in determining attrition, especially for girls, and we were forced to abandon this approach and rely on the provided weights.

ates; 4) both parents have a post-secondary education below the BA level or one has post-secondary education below the BA level and the other has a lower level of education; 5) one parent has a BA and the other has some lower level of education; and 6) both parents have at least a BA. Lone parent families are assigned to categories 1, 3, 4 or 6, depending on the parent's education. The sample means in Table A.1 indicate that approximately 10 percent of the sample falls in each of categories 1 and 6. We also include a set of variables reflecting family structure, with indicators corresponding to lone parent families, two parent families in which both biological parents are present, and "other" two parent families which correspond, essentially, to step-parent families, and other family types (the omitted category is a two biological parent family). Lone parent families may face a "poverty of time" that implies stresses that affect school completion. We include a dummy variable for whether the person lives in a rural (as opposed to urban) location and a variable corresponding to the number of times the family has moved in the child's lifetime up to age 15. We would expect more moves to correspond to a weakening of social connections that may be important in school completion. Finally, we also include variables corresponding to whether the child is an immigrant and whether the youth is of aboriginal descent.[28] These variables are included because of evidence that recent immigrants are facing substantial barriers to integrating into the economy and society at large. The aboriginal descent variable is suggested by high rates of poverty in this community.

We present all of our results separately by gender. In the first column of Table 2.1, we present, for males, the marginal effect of parental education and family income on the probability of dropping out of high school. The estimates correspond to the impact on the probability of dropping out of increasing adult-equivalent income for a family of 4 from \$15,000 to \$50,000 and of changing from the omitted parental education category (in which both parents have a BA or higher education) to the listed education category. The marginal effects are evaluated at mean income and the omitted parental education category. In the top panel of the ensuing tables,

[28]In specifications not shown, we included indicators for whether the child is second generation (i.e., born in Canada with at least one parent who is an immigrant) and the language spoken at home is an official language. Because these variables were never significant or economically substantial in a variety of specifications we dropped them from the analysis

we present the probability of dropping out for a youth whose parents both have a high school diploma, evaluated at the mean of the other variables. We present the predicted probability of dropout by parental education in Figure 2.1 for the specifications in columns 2 and 3 to aid the reader in interpreting these numbers.

The first point of interest from the estimates is that the association between family income and dropping out for boys is weak. An increase in income per adult equivalent for a family of 4 from $15,000 to $50,000 reduces the probability of dropping out by less than .01. This fits with results in Belley et al. (2008) indicating that while there are family income effects on educational attainment in Canada, they are not strong. We view parental education as related to permanent income for the family, therefore current income when controlling for parental education is something closer to transitory income. In specifications where we do not control for parental education, the coefficient on family income is twice as large.

The remaining entries in column 1 show that parental education is very strongly correlated with dropping out. Relative to a student both of whose parents have a BA or higher education (a person whose probability of dropping out is .007), a student both of whose parents are themselves high school dropouts has a .15 higher probability of dropping out. Youth whose parents have a high school diploma have a .05 higher probability of dropping out compared to those whose parents both have a BA. The main conclusion from the first column is that there is a steep gradient associated with parental education which points toward a calcification of educational differences across generations. Belley et al. (2008) show that dropout gradients with respect to parental education and family income are steeper in the US, but the evidence in this table indicates that inter-generational persistence is still an issue in Canada.

Column 1 of Table 2.2 contains the results from the same specification for the female sample. The marginal association with family income is even smaller and is not statistically significant. Dropout rates are much lower for girls (e.g., the probability of dropping out when both parents are dropouts is .16 for boys and .08 for girls) but the family gradient is still significant.

In column 2 of Table 2.1 (for boys), we expand the definition of socio-economic status to include family structure, number of siblings, age in months, residence in a rural location, immigrant status and aboriginal status. In the top panel of Figure

2.1, we plot the probabilities associated with different parental education levels from this specification.[29] This figure documents the substantial size of the parental education gradient that is present even when we control for other socio-economic variables. Those other variables also have substantive effects in their own right. Thus, being in an "other" two parent family is associated with a sizeable increase in the probability of dropping out (by .08) relative to a family with two biological parents present. On the other hand, the association between dropping out and being from a lone parent family is relatively small and statistically insignificant. Being of Aboriginal descent is associated with a .058 higher probability of dropping out, all else equal. Since all else is not equal, the actual average difference in dropping out between Aboriginals and other Canadians will be much larger than this.

The results for females in column 2 of Table 2.2 and plotted in the top right graph in Figure 2.1 again show smaller dropout probabilities and a flatter parental education gradient. The results for other socio-economic variables are similar to those for boys but generally smaller in magnitude. As for boys, including these controls reduces the size of the parental education gradient, but not to a large degree.

The end result from this initial look at the data is that several dimensions of socio-economic status have strong associations with dropping out (aboriginal status and "other" family type in particular) but that the strongest association is with parental education. This association indicates a substantial socio-economic gradient and is our main point of focus.

2.5.2 Proxy Estimator

Next we turn to including proxy variables for child's cognitive ability and parental valuation of education in our Probit specification. We wait until later to introduce non-cognitive ability proxies in order to match much of the previous literature (e.g. Todd and Wolpin 2006). We proxy for cognitive ability using dummy variables corresponding to the student being in quartiles 1, 2 and 3 of the distribution of reading

[29] In this figure, each grouping of bars shows the probability of dropping out for a base person (i.e., a person with the average family income and average values of the other socio- economic variables) for varying levels of parents' education (specified under the individual bars). The education categories reported in Figure 2.1 correspond to families where both parents have the specified education level.

scores from the PISA test (with individuals in the top quartile being the omitted category). It is worth re-iterating that in our model the PISA score is a sufficient statistic for everything affecting cognitive ability before age 15 and that controlling for it changes the interpretation of the socio-economic variable effects to the impacts of these characteristics on changes in education outcomes relative to the age 15 baseline established by the PISA score. To proxy for parental aspirations, we use parental responses to the question, "What is the highest level of education that you hope [child's name] will get?" We code a dummy variable equalling one if the parent's response was a 'university degree or higher' and zero if their response corresponded to a 'college degree or lower'.[30] We argued in Section 3 that including these proxies should help reduce the extent to which socio-economic gradient coefficients are capturing ability and parental valuation factors but may not eliminate those problems completely.

We introduce the PISA quartiles interacted with parental aspirations in column 3 of Tables 2.1 and 2.2. We also plot the predicted probabilities of dropping out (evaluated at the mean) for the various categories in Figure 2.2. These patterns are interesting in themselves. When boys scored in the top quartile on the PISA reading test they were very unlikely to dropout. For these boys, changing parental aspirations is associated with a small and statistically insignificant change in the dropout probability. Boys who scored in the top PISA quartile and whose parents hoped they would achieve a university degree had less than a one percent chance (0.008) of dropping out, compared to a 3 per cent chance for similar boys whose parents expected a lower level of educational attainment.

In the bottom three PISA quartiles, parental aspirations have significant impacts that increase in magnitude as we move lower in the PISA distribution. Figure 2.2 indicates that high parental aspirations not only reduce the likelihood of dropping out, but also flatten the gradient across the PISA quartiles. This happens in a non-linear way: the largest proportional reductions in dropping out occur for students in the 2nd and 3rd quartiles.

One way to put these results in context is to consider where, along the PISA distribution, the probability of dropping out is closest to the unconditional proba-

[30]Some parents responded 'Any level above high school'. These responses were coded as 0.

bility and how that differs by parental expectations. Overall, in the sample .055 of boys drop out. Boys whose parents have low aspirations will drop out at the unconditional average rate only when they have reading scores in the third quartile. If a boy's parents expected him to obtain a university degree, his chances of dropping out are similar to the average if he is in the bottom quartile of the PISA distribution. The overall implication is that high parental aspirations have a powerful influence on educational outcomes.

For girls, the same patterns are evident but at lower probability levels. In addition, parental aspirations appear to have a much smaller effect for girls. While reducing aspirations from low to high cuts dropout probabilities by more than one half for boys at all ability levels, for girls the reduction is much smaller. For the second ability quartile girls, the reduction is only from .033 to .025.

In the bottom panel of Figure 2.1, we show the impacts of parental education when controlling for the set of interacted parental aspirations and PISA scores. Comparing this to the top panel of Figure 2.1 shows the impact of controlling for these variables. For boys, the difference in the dropout probability between those with two BA parents and those with two dropout parents is .13 when just controlling for income, parental education and other socio-economic variables, but falls to .078 when also controlling for aspirations and PISA scores. For girls, once we control for aspirations and PISA score, the impact of parental education is reduced by similar proportions to those for boys. Thus, a substantial proportion of the parental education effects we estimated in the earlier specifications are actually masked aspiration and ability effects. A girl with both parents highly educated and one with both parents who are dropouts differ in their probability of dropping out by only .03 once we control for aspirations and ability. Controlling for these factors also reduces the associations with family status and aboriginal status.

In the final two columns of Tables 2.1 and 2.2, we introduce PISA and parental aspirations separately to investigate whether one of these variables is more important in reducing the socio-economic gradient. When we include only the parental aspirations in column 4, we find that boys whose parents are high school dropouts have a .105 higher probability of dropping out relative to boys whose parents both have a BA. This suggests that parental aspirations alone can account for roughly half the reduction in the socio-economic gradient obtained from including aspirations and ability together. For girls, parental aspirations appear to have little impact on the socio-economic gradient. Controlling for expectations only narrows the 2 BA versus 2 dropout parent difference by .009. When we control for PISA alone, in column 5, for boys the marginal impact of having parents who dropped out of high school, relative to parents with a BA, is .09. Because PISA and parental aspirations are highly correlated it is difficult to reach a conclusion about the relative importance of these measures using such reduced form results. As outlined in our model, we are not directly interested in PISA and the stated aspirations of parents but in the underlying factors of ability and parental valuations of education. In subsequent sections, we move to a factor model which can identify these unobserved heterogeneity terms.

Controlling for non-cognitive abilities, behavioral choices and school characteristics

Before moving to the estimation of unobserved heterogeneity, we consider whether the general patterns observed in Tables 2.1 and 2.2 persist in the proxy set-up after we control for non-cognitive skills, some behavioral choices and school characteristics. For comparison, the first column of Tables 2.3 and 2.4 reproduce column 3 of Tables 2.1 and 2.2 respectively. In column 2, we introduce the proxy measure for non-cognitive skills that we described earlier, and which takes on the value one if a child said that he never wanted to 'just get by'. This variable, which measures conscientiousness, is significantly related to dropping out for both boys and girls in the statistical sense, but the effect is small in magnitude. Never wanting to just get by reduces the probability of dropping out by .024 for boys and .012 for girls. Moreover, for both boys and girls, the socio-economic gradient, as well as the gradients associated with PISA and parental aspirations change very little after including the proxy for non-cognitive skills.

In the next column of Tables 2.3 and 2.4, we add our second measure of non-cognitive skills (an indicator variable equalling one if the student reports he always completes his assignments) and two scale measures of non-cognitive skills, self-esteem and self-efficacy. Self-esteem is measured using the 10-item Rosenberg's self-esteem scale and captures the youths' global feelings of self-worth or self-acceptance (see Rosenberg 1965). Because this measures overall psychological well-being, we anticipate that its relationship to behavioral outcomes may be weak. The YITS includes a self-efficacy scale adapted from Pintrich and Groot (1990) which measures perceived competence and confidence in academic performance. The self-efficacy scale has a mean of .10 and standard deviation of 1.4.

The results in column 3 indicate that self-esteem has no direct effect on dropping out for either boys or girls, but self-efficacy has a significant impact on boys, although a relatively small one. A one standard deviation increase in self-efficacy reduces the probability of dropping out by .01. For girls, the effect is virtually zero and is statistically insignificant. The third column also shows that children who complete their assignments are less likely to dropout by a margin of .024 for boys and .01 for girls.

Inclusion of self-efficacy, self-esteem and the homework indicator does not affect the socio-economic gradient but does reduce the impact of parental aspirations and PISA scores. As we mentioned earlier, the self-efficacy scale likely also captures cognitive ability. For example, one question included in the scale asks students to indicate how frequently this statements is true: 'I'm confident I can understand the most complex material presented by the teacher'. This, along with the correlation between parents' aspiration and PISA, would explain why including self-efficacy in column 3 reduces the PISA-aspirations gradient. Nonetheless, the PISA and aspirations gradients remain strong.

In column 4, we include variables corresponding to choices made by the students. One key choice a person makes is their group of friends. Many papers emphasize the potential impact of peers on schooling and other outcomes. We generate a series of indicator variables corresponding to whether all of the respondent's close friends: think completing high school is very important; skip classes once a week or more; have dropped out of high school; are planning to attend post-secondary education; have a reputation for causing trouble; smoke cigarettes; think it's okay to work hard at school. Thus, these variables capture a combination of risky behaviours and attitudes toward schooling among friends. We use the extreme versions of these questions (i.e., that all versus just some friends take these attitudes and behaviours) to give these peer variables their maximum possible impact. Even with this, the results indicate mixed evidence on potential peer effects. For boys, only the variable corresponding to having friends who are planning to get more education after high school is statistically significant and its effect is not substantial. Similarly, for girls, the peer variables tend not to be statistically significant and do not imply economically substantial effects.

We also include a dummy variable for whether the youth has a dependent child, since US studies indicate that this may correlate with dropping out, particularly for girls. In fact, we find that this variable has a positive effect on dropping out for both boys and girls. In interpreting the estimated coefficient for boys it is worth noting that respondents are not asked if they have ever fathered a child but whether they have a dependent child. Thus, this result says that boys who have both fathered a child and have taken the responsibility to help raise it are very likely to have dropped out - possibly in order to get a job to support their young family.

Finally, we include an indicator variable for whether the individual smoked at least weekly at age 15. Smoking is often seen as a marker of risky behavior in general and is sometimes interpreted as reflecting youth with relatively high discount rates. Thus, one would expect youth who smoke to drop out more often than those who do not since, under this interpretation, they do not value the future as highly. For boys, smoking turns out to be a strong predictor of dropping out, raising the probability of dropping out by nearly .052. For girls, it is also statistically significant but has a much smaller effect.

Introducing the peer, dependent child and smoking variables has very little impact on the socio-economic gradient impact estimates but does generate a reduction in the size of the aspiration/PISA effects. This indicates that ability and parental aspirations are correlated with risky behaviours such as smoking and sexual activity. We recognize that selection and endogeneity concerns complicate the interpretation of the coefficients on the set of variables introduced in column 4 and so do not include them in the remaining specifications. But whatever their own effects, their inclusion does not alter our main conclusions about the socio-economic gradient and its relationship to aspirations and ability. [31]

In the final column of Tables 2.3, we incorporate school characteristics that

[31] We also estimated specifications in which we included measures of hours of paid work for the students. None of the measures we examined entered significantly or changed our key estimated marginal impacts. Working is very common among Canadian youth, both during the school year and during the summer.

were reported by the high school administrators as a part of the first wave of the YITS survey. We include a dummy variable that takes a value of one if the administrator reported that a lack of instructional material hindered the learning of grade 10 students to some extent or a lot. We also include the ratios of students to teachers, and students to computers. We include these measures of school characteristics as a way to gauge the sensitivity of our results. Because we have not addressed the endogenous selection of families into schools, one should not interpret these results as causal. With that caveat in mind, the results in the fifth column do indicate that the student to teacher ratio in a school is correlated with dropping out for boys. Including school characteristics has essentially no effect on the socio-economic gradient and little impact on the PISA and parental aspirations effects.[32]

The last column of Table 2.3 also includes the local youth unemployment rates, which reflect the relative attractiveness of the alternatives to school. Youth unemployment in this reduced form set-up does not appear to be correlated with dropping out.

We could not include school characteristics in our reduced form estimates for girls because several of the marginal effects could not be identified. For at least two reasons there is not enough variation in the data to include school characteristics. The first reason is simply that very few girls drop out and variables such as parental education and PISA explain virtually all of the variation. The second reason stems from residential sorting which means that different types of families are clustered together with their children attending the same schools. Practically, this means that for a large fraction of the schools in our data, no girls dropped out.

Given the findings of this section, in the remainder of the paper we examine the role of abilities and parental aspirations in determining the socio-economic gradient without considering peers or school effects. This allows for a sharper focus on a set of relationships that, anyway, appear to be little or not affected by these effects. We also restrict our attention to boys since they face a much more substantial risk of dropping out. Preliminary investigations with our more complicated estimators indicated that there was not enough variation in dropping out among girls to support a deeper investigation.

[32]Non-response to the administrators survey significantly reduces the sample, as a result we do not include any of these school measures in the unobserved-factor models we present later.

2.5.3 Factor Estimators

In this section, we present results from the full factor models set out in Section 2.3. Recall that our goal with the basic factor model is to use the added measures of ability (cognitive and non-cognitive) and parental valuations available in the YITS to better control for these factors. As we stated earlier, if the relationships among factors, measurement equations and the behavioral equation are all linear (as is typically assumed when implementing these systems) then our basic system estimator provides consistent estimates of all the parameters in the model.

A key decision in implementing these models is the number of points of support in the estimated factor distributions. We first estimated the models with two points of support for each factor then added additional points. Adding a third point of support for the cognitive ability factor distribution significantly improved the fit of the model but adding a third point for the parental valuation and non-cognitive ability distributions, as well as a fourth point of support for cognitive ability, were not helpful.[33] Thus, we implement a specification with three points of support for cognitive ability and two each for parental valuation and non-cognitive ability. Table 2.5 reports the marginal impact of socio-economic background on the probability of dropping out estimated from our two different factor models. In the first column, we present results from the basic model where the probabilities associated with the mass points for the three factors do not vary with parental education. The results in this first column are directly comparable to those presented in column 3 of Table 2.1 for the proxy estimator.

Once again, the impact of family income is essentially zero: notably smaller than the already small impacts in Belley et al. (2008). Since they control for test scores and grades but not parental education valuation, the inclusion of the latter effects appears to account for the complete lack of family income effects in our estimation. As with the proxy estimator results, the remaining socio-economic variables all have statistically significant and economically substantial effects. The estimated effects from our basic factor model are very similar in magnitude to the effects from our proxy model. For example, the difference in dropout rates between a teenager from a two BA family and a teenager from a two dropout family is .078

[33] Specifically, the model returned probability masses for the additional points of support that were very close to zero and imprecisely estimated.

in the proxy model and .075 in the factor model. This implies that the bias related to the proxy model, discussed in Section 3, appears to be small, which could arise if the variance of the measurement error in (2.11) is relatively small.

Estimates of the factor loads, locations and associated probabilities are given in Table A.3 and Table A.4. The factor loads indicate that all three factors have statistically significant and sizeable effects on both the dropout and grades indexes. Interestingly, the child non-cognitive ability factor does not have a significant effect in either measurement equation related to parental valuation ($parpref$ and *saved*). The cognitive ability factor enters both equations significantly but the impact in the *saved* equation is small. Thus, at least for abilities we can measure, our measures of parental responses and actions relating to their valuation of education are not simply reflections of the child's abilities. To the extent this carries over to any other, unmeasured factors, this pattern implies that we really are capturing parental valuation of education rather than getting another measure of child abilities. Finally, the parental valuation factor is a significant determinant of the non-cognitive ability measures. Parents who value education induce their children to complete their assignments on time.

In Table 2.6, we describe the joint impacts of parental education, ability and parental valuation of education by presenting fitted probabilities for each possible combination of the factors and parental education. (The predicted probabilities shown in Tables 2.6 and 2.7 are also shown in Figures 2.3 and 2.4, respectively.) In particular, for a given level of parental education, we form a fitted probability by setting all other variables at their mean values and then using the estimated mass point location value for one of the two points in the parental valuation and non-cognitive factor distributions and one of the three points in the cognitive ability distribution. This yields 12 fitted probabilities for each education level. The top panel of Table 2.6 shows the fitted dropout probabilities evaluated at high non-cognitive ability and the bottom panel shows the same for low non-cognitive ability.

Several points emerge from this exercise. First, having high cognitive ability at age 15 implies an almost zero probability of dropping out regardless of non-cognitive ability, parental educational valuation or parental education. The one exception to this is for teenagers with low non-cognitive skills whose parents have low educational valuation and both of whom are high school dropouts. Those teenagers have a predicted probability of dropping out of .017, which is still below the overall average. In contrast, low ability teenagers have substantial probabilities of dropping out, particularly when their parents place a low value on education. Thus, ability measured at age 15, and through it all the factors earlier in life that helped to shape it, is a very important determinant of dropping out.

Second, parental valuation of education has sizeable effects for teenagers with medium and low ability. For example, a teenager with low ability in both domains and both of whose parents are high school graduates has a probability of dropping out of .08 if his parents value education highly but .29 if they have a low valuation. Proportionally, the importance of parental valuation is largest for the medium ability group, though in absolute terms it is even larger for the low ability group. Considering just those with low non-cognitive ability, medium cognitive ability teenagers whose parents value education highly look very much like high ability teenagers from low valuation families and, like them, are very unlikely to drop out. Medium ability teenagers with low valuation parents, in contrast, have large probabilities of dropping out.

Third, having high non-cognitive skills reduces the chances of dropping out and can offset the impact of coming from a family with a low valuation of education. However, the impact of non-cognitive skills is modest relative to the effect of parents' valuations. For a mid-cognitive and low-non-cognitive ability youth whose parents are dropouts and have a low valuation of education, raising his non-cognitive ability reduces the chance that he will drop out by 6 percentage points (roughly one third). In contrast, changing his parents' valuation reduces the probability of dropping out by 12 percentage points, a reduction of 80 percent.

Finally, there remains a significant gradient with respect to parental education even holding ability and parental valuation of education constant. For example, the difference in the dropout probability between a teenager with low cognitive and non-cognitive abilities from a high valuation family both of whose parents are dropouts and the same teenager with two BA parents is .15. The implication is that holding constant ability at age 15 and all the factors determining it, higher educated parents still impart something more to their children.

In Section 2.3, we argued that the standard linear system of equations in a factor model may not be an accurate depiction of the ability generation and school attainment process. We allow for greater flexibility in our extended system estimator by allowing the probabilities associated with the points of support for the unobserved factors to differ by parental education level. Note that this specification nests the more standard model as a special case. A likelihood ratio test rejects the restrictions that the factor probabilities not differ by parental education at any conventional significance level.[34]

The estimated marginal impacts from the extended factor model are given in the second column of Table 2.5. Allowing the factor distributions to differ by parental education reduces the impacts of all the socio-economic variables to some extent. Most importantly, it effectively eliminates the gradient with respect to parental education. The difference in the probability of dropping out relative to a boy from a BA-family is not statistically significant for any of the other parental education categories. The effect size is less than one percentage point for youth whose parents have a high school diploma. Under the single-index assumption discussed earlier, these results imply that, while parental education may influence the teenager's level of ability up to age 15, parents with higher levels of education do not impart anything more to their children after age 15 once we condition on their valuation of education.

In Table 2.7, we show the fitted probabilities of dropping out by factor-type and parental education for the extended factor model. As in the simpler factor model, ability has substantial effects on dropping out.[35] With few exceptions, high cognitive-ability teenagers do not drop out regardless of their parent's education or the values of the other factors. For teenagers with low non-cognitive skills whose parents are high school dropouts and place low value on education, moving from high to low cognitive ability increases the probability of dropping out to .36. But parental valuation effects are nearly as large. A teenager with low cognitive and non-cognitive abilities whose parents place a high value on education has about a .03 probability of dropping out which means the impact of parents' valuations for a low-ability boy is .33. Moreover, a student whose parents place a high value on

[34]Specifically, the test statistic is distributed as $\chi^2(20)$ and takes a value of 977.59.

[35]Note that the small and insignificant differences in marginal impacts of parental education in Table 2.5 are converted into larger (though still statistically insignificant) differences in Table 2.7 because the curvature of the normal cumulative distribution function toward the tails converts small parameter differences into larger probability differences.

education has essentially a zero probability of dropping out unless he has both low cognitive and non-cognitive abilities, and even then his dropout probabilities are very low. In comparison, non-cognitive ability has effects that are substantial but less than either of the other two factors. Looking at the bottom right corner of the panels, increasing from low to high non-cognitive ability reduces the probability of dropping out by .17. This compares to reductions on the order of .3 from improving cognitive and parental valuation from the same starting point.

Table 2.8 shows the estimated distribution of teenagers across the points of support for each of the factors and each parental education level. Table 2.9 shows the joint distributions. These tables show that teenagers whose parents both have a BA have high probability (.44) of having high cognitive ability and having parents who place a high value on education. In contrast, .41 of teenagers from two dropout families are in the low cognitive ability - low parental valuation category and a further .24 are in the medium cognitive ability - low valuation category. Thus, raw comparisons of teenagers from these two types of families will capture the fact that the children in the two BA families are more able and have parents who care more about education. The differences in ability reflect differences up to age 15. These could include inter-generational transmission of ability (equation 2.1) but could also include the type of early childhood investment effects stressed in recent papers by Heckman and co-authors (e.g. Heckman and Lochner 2000) (equation 2.3). Interestingly, non-cognitive ability does not show the same degree of correlation with parental education. In fact, teenagers whose parents both have a BA are less likely to have high non-cognitive skills. Roughly one third of the boys from BA families have high non-cognitive skills compared to about 46 percent of the boys whose parents are both dropouts.

It is difficult to summarize the impact of parents' valuations because the impacts are very different at different points of the ability distribution. However, we can gain some insight through a counterfactual experiment in which we predict the dropout probability for a boy whose parents are dropouts, attributing to him the valuation distribution of a BA-family and holding constant the distribution of his abilities. This counterfactual probability is only .037 compared to the actual predicted probability of .13 for a boy in a two dropout family.

An interesting implication of these results is that it is not parental education that matters for dropping out but parental valuation of education. That is, a child whose parents are both themselves dropouts has the same probability of dropping out as a child with the same ability from a highly educated family if his parents care as much about education as their more educated counterparts. In our data, what parents who care about education actually impart is a black box - they could devote more resources to their child's education, they might convince their child that there is a return to effort in school, or they might enforce effort of at least some minimal level. The fact that parental education does not matter points away from resource based arguments since we would expect parental education to be correlated with family permanent income. Our results are in apparent agreement with Behrman and Rosenzweig (2002)'s estimates of the impact of variation in maternal education on children's educational outcomes holding constant family fixed factors (by using differences in education between twin mothers). But they find significant effects of paternal education. Carneiro et al. (2007) find significant effects of maternal education on grade retention in NLSY data when they instrument for maternal education using local labour market conditions and access to colleges when the mother was 17. To the extent that increasing parental education increases parental aspirations for their children's education (which Carneiro et al. (2007) find is the case), the estimated parental education effects identified by these papers may ultimately occur through the channel we identify.

2.6 Conclusion

This paper attempts to provide some insights into the strong correlation between parental education levels and the educational outcomes of their children. Such a correlation suggests a calcification of educational inequalities that may, in turn, result in social efficiency losses and/or lack of fairness of opportunities. Understanding the source of the correlation is, therefore, a necessary first step in deciding whether policy interventions are called for and, if so, what form those interventions should take. The key empirical challenge is to decide whether the correlation reflects direct effects of parental education (perhaps because more educated parents are better able to help their children with their school work) or is actually capturing the influence of sources of heterogeneity that are typically unobservable, such as cognitive ability. Combining insights from earlier work such as Sewell et al. (1969), Davies and Kandel (1981), Todd and Wolpin (2006) and Cunha and Heckman (2007; 2008), we investigate the importance of three underlying factors in determining the propensity of teenagers to drop out of high school: cognitive abilities, non-cognitive abilities, and the value placed on education by the teenager's parents.

Our analysis of the unobserved factors is motivated by some compelling reduced form evidence and is based on a simple life-cycle model of education choice. Using Canadian micro-data we establish the size of the so-called socio-economic gradient of dropping out, which relates schooling decision to underlying family factors. As might be expected, we find that less educated family backgrounds are associated with a higher incidence of dropping out. Next, by exploiting finer information available in the YITS, we find that cognitive ability and parental valuations of education play a major role in determining the family effect, especially for boys. When we control for these factors through proxies, the socio-economic gradient is largely reduced, in many cases by up to a half. We take this as indication that family background differences may reflect differences in cognitive skills and family valuations of education.

We investigate this finding in more detail using an unobserved factor approach based on Carneiro et al. (2003). Given arguments in Cunha et al. (2006) that ability production functions may be non-linear in parental investments, we consider an extension to the estimator in Carneiro et al. (2003) in which the unobserved factor distributions are allowed to differ across families with different parental education, both in terms of shape and location. Implementing this unconstrained estimator results in four main findings. First, ability at age 15 has a substantial impact on dropping out. The highest ability individuals are predicted never to drop out regardless of parental education or parental valuation of education. In contrast, the lowest ability teenagers have a probability of dropping out of approximately .36 if their parents have a low valuation of education. Second, parental valuation of education has a substantial impact on medium and low ability teenagers. A low ability boy has a probability of dropping out of approximately .03 if his parents place a high value on education but .36 if their educational valuation is low. That parental concerns about education have an impact on educational outcomes will not seem surprising to anyone who has stood over their teenage son to make him do his homework. Third, non-cognitive ability has impacts that are sizeable but much smaller than those of the other two factors. Fourth, parental education has no direct effect on dropping out once we control for ability and parental valuation of education.

The interpretation of these results depends on the underlying economic structure. If we assume an index-type model in which cognitive and non-cognitive abilities at age 15 fully summarize all inputs before that age which are relevant for subsequent dropout decisions then, once we include measures of those abilities, all other effects should be interpreted as being impacts of the relevant factors and characteristics after age 15. Thus, our results would indicate that both parental education and value put on education may influence abilities at age 15 (perhaps through impacts in early childhood) but only parental valuation of education has an effect beyond this age.

Whether or not one accepts the assumption that cognitive and non-cognitive abilities at age 15 are sufficient statistics for everything that has gone before, two striking results remain: first, only parental valuation of education (rather than parental education itself) matters in the dropout decision, once we control for abilities; second, the quantitative impact of parental valuation is very large. We view these results as hopeful for policy, since parental valuations of education are potentially amenable to interventions which do not require very long time frames, unlike targeting parental education itself. The set of policies that seem interesting given our results are ones which either coach the parents themselves (such as the Baby College program) or find ways to replicate what high valuation parents do for their children (such as expanded Big Brothers and mentoring programs, or extended hours in school or publicly provided care). Whichever way one interprets our results, parental valuation of education falls into the category of determinants of education (and through it, outcomes in later life) for which a youth cannot be held morally responsible and this provides some justification for policy intervention.

Table 2.1: Factors affecting dropping out of high school among boys Marginal effects estimated in a Probit predicting dropping out at age 19. (Standard errors in parenthesis)

	1	2	3	4	5
Predicted probability of dropping out for reference person					
	0.089	0.054	0.037	0.043	0.044
Log family income	-0.004 (0.002)**	-0.002 (0.001)	-0.003 (0.001)***	-0.003 (0.001)**	-0.002 (0.001)*
Parents' highest educational attainment –Reference both parents have a BA or higher					
One parent has BA	0.023 (0.009)***	0.019 (0.008)**	0.015 (0.008)*	0.018 (0.008)**	0.017 (0.008)**
At least one parent has PSE below BA	0.049 (0.007)***	0.040 (0.007)***	0.024 (0.007)***	0.030 (0.007)***	0.029 (0.007)***
Both parents have a high school diploma	0.050 (0.012)***	0.047 (0.012)***	0.028 (0.011)***	0.035 (0.012)***	0.035 (0.011)***
One parent has a H.S. diploma	0.076 (0.016)***	0.063 (0.015)***	0.032 (0.011)***	0.045 (0.013)***	0.040 (0.012)***
Both parents have less than H.S.	0.154 (0.023)***	0.133 (0.022)***	0.078 (0.017)***	0.105 (0.020)***	0.090 (0.017)***
PISA scores and parents' aspirations– Reference PISA Quartile 4 and BA aspirations					
Below BA aspirations–PISA Quartile 1			0.137 (0.026)***		
Below BA aspirations–PISA Quartile 2			0.073 (0.020)***		
Below BA aspirations–PISA Quartile 3			0.045 (0.018)**		
Below BA aspirations–PISA Quartile 4			0.022 (0.015)		
BA and above aspirations–PISA Quartile 1			0.055 (0.016)***		
BA and above aspirations–PISA Quartile 2			0.018 (0.007)**		
BA and above aspirations–PISA Quartile 3			0.011 (0.007)*		
Parents' aspirations –Reference below BA aspirations					
BA and above aspirations				-0.068 (0.015)***	
PISA reading scores –Reference PISA Quartile 2					
Quartile 1					0.099 (0.019)***
Quartile 2					0.038 (0.012)***
Quartile 3					0.017 (0.008)**
Province dummies	N	Y	Y	Y	Y
Family background controls	N	Y	Y	Y	Y
Sample size	7,755	7,755	7,755	7,755	7,755

continued, next page

Table 2.1: Factors affecting dropping out of high school among boys (cont'd) Marginal effects estimated in a Probit predicting dropping out at age 19. (Standard errors in parenthesis)

	1	2	3	4	5
Predicted probability of dropping out for reference person					
	0.089	0.054	0.037	0.043	0.044
Other family characteristics					
Aboriginal		0.058 (0.030)*	0.028 (0.021)	0.035 (0.025)	0.034 (0.023)
Immigrant		-0.027 (0.014)*	-0.019 (0.011)*	-0.014 (0.014)	-0.028 (0.011)**
Rural		-0.010 (0.009)	-0.016 (0.007)**	-0.016 (0.008)**	-0.013 (0.007)*
Number of moves		0.003 (0.002)**	0.003 (0.001)**	0.003 (0.001)**	0.003 (0.001)**
Age in months		0.003 (0.002)**	0.003 (0.001)**	0.003 (0.001)**	0.003 (0.001)**
Number of siblings		-0.006 (0.004)	-0.003 (0.003)	-0.004 (0.004)	-0.004 (0.004)
Family structure –Reference two biological parent families					
Other two parent families		0.080 (0.022)***	0.052 (0.016)***	0.063 (0.019)***	0.061 (0.018)***
Lone parent family		0.008 (0.012)	0.012 (0.010)	0.011 (0.011)	0.011 (0.011)
Control for PISA scores	N	Y	Y	N	Y
Control for parental aspirations	N	N	Y	Y	N
Province dummies	N	N	Y	Y	Y
Sample size	7,755	7,755	7,755	7,755	7,755

Source: Youth in Transition Survey, Cycle 3 (Cohort A)
Estimates are weighted to account for non-response to the parents' survey and longitudinal attrition.
Standard errors clustered by high school.
*** indicates result is statistically significant at .01 level, ** at .05 level, * at .10 level
Marginal effects are evaluated for a youth whose parents both have a high school diploma at the mean of all the other variables.

Table 2.2: Factors affecting dropping out of high school among girls. Marginal effects estimated in a Probit predicting dropping out at age 19. (Standard errors in parenthesis)

	1	2	3	4	5
Predicted probability of dropping out for reference person					
	0.054	0.021	0.014	0.018	0.016
Log family income	-0.0002 (0.0002)	-0.0001 (0.0001)	-0.0001 (0.0001)	-0.0001 (0.0001)	-0.0001 (0.0001)
Parents' highest educational attainment –Reference both parents have a BA or higher					
One parent has BA	0.006 (0.004)	0.005 (0.005)	0.003 (0.006)	0.004 (0.005)	0.004 (0.006)
At least one parent has PSE below BA	0.027 (0.006)***	0.021 (0.006)***	0.012 (0.006)*	0.018 (0.006)***	0.015 (0.006)**
Both parents have a high school diploma	0.021 (0.008)***	0.017 (0.008)**	0.009 (0.007)	0.014 (0.007)*	0.012 (0.007)
One parent has a H.S. diploma	0.070 (0.016)***	0.052 (0.012)***	0.027 (0.009)***	0.043 (0.012)***	0.032 (0.010)***
Both parents have less than H.S.	0.075 (0.014)***	0.060 (0.013)***	0.030 (0.010)***	0.051 (0.012)***	0.034 (0.010)***
PISA scores and parents' aspirations– Reference PISA Quartile 4 and BA aspirations					
Below BA aspirations–PISA Quartile 1			0.065 (0.019)***		
Below BA aspirations–PISA Quartile 2			0.027 (0.012)**		
Below BA aspirations–PISA Quartile 3			0.023 (0.011)**		
Below BA aspirations–PISA Quartile 4			0.007 (0.007)		
BA and above aspirations–PISA Quartile 1			0.045 (0.016)***		
BA and above aspirations–PISA Quartile 2			0.020 (0.008)**		
BA and above aspirations–PISA Quartile 3			0.000 (0.002)		
Parents' aspirations –Reference below BA aspirations					
BA and above aspirations				-0.022 (0.008)***	
PISA reading scores –Reference PISA Quartile 2					
Quartile 1					0.054 (0.015)***
Quartile 2					0.021 (0.008)***
Quartile 3					0.005 (0.004)
Province dummies	N	Y	Y	Y	Y
Family background controls	N	Y	Y	Y	Y
Sample size	8,376	8,376	8,376	8,376	8,376

continued, next page

Table 2.2: Factors affecting dropping out of high school among girls (cont'd) Marginal effects estimated in a Probit predicting the dropping out at age 19. (Standard errors in parenthesis)

	1	2	3	4	5
Predicted probability of dropping out					
	0.054	0.021	0.014	0.018	0.016
Other family background characteristics					
Aboriginal		0.030	0.022	0.022	0.021
		(0.020)	(0.015)	(0.014)*	(0.016)
Immigrant		-0.020	-0.015	-0.017	-0.017
		(0.007)***	(0.005)***	(0.006)***	(0.006)***
Rural		-0.001	-0.002	-0.003	-0.002
		(0.005)	(0.004)	(0.005)	(0.004)
Number of moves		0.002	0.001	0.002	0.002
		(0.001)**	(0.001)**	(0.001)**	(0.001)**
Age in months		0.002	0.001	0.002	0.002
		(0.001)**	(0.001)**	(0.001)**	(0.001)**
Number of siblings		-0.004	-0.003	-0.003	-0.003
		(0.002)*	(0.002)*	(0.002)	(0.002)*
Family structure –Reference two biological parent families					
Other two parent families		0.028	0.016	0.022	0.019
		(0.012)**	(0.008)*	(0.010)**	(0.009)**
Lone parent family		0.009	0.009	0.009	0.010
		(0.008)	(0.007)	(0.008)	(0.008)
Control for PISA scores	N	N	Y	N	Y
Control for parental aspirations	N	N	Y	Y	N
Province dummies	N	Y	Y	Y	Y
Sample size	8,376	8,376	8,376	8,376	8,376

Source: Youth in Transition Survey, Cycle 3 (Cohort A)
Estimates are weighted to account for non-response to the parents' survey and longitudinal attrition.
Standard errors clustered by high school.
*** indicates result is statistically significant at .01 level, ** at .05 level, * at .10 level
Marginal effects are evaluated for a youth whose parents both have a high school diploma at the mean of all the other variables.

Table 2.3: The effects of school characteristics, and non-cognitive and behavioral factors on dropping out for boys Marginal effects estimated in a Probit predicting dropping out at age 19. (Standard errors in parenthesis)

	1	2	3	4	5
Predicted probability of dropping out					
	0.037	0.035	0.030	0.029	0.023
Log family income	-0.003 (0.001)***	-0.002 (0.001)**	-0.002 (0.001)	-0.002 (0.001)**	-0.001 (0.001)
*For a family of four, marginal effect of a difference in income from $50,000 and $1,000					
Parents' highest educational attainment –Reference both parents have a BA or higher					
One parent has BA	0.015 (0.008)*	0.015 (0.008)*	0.011 (0.006)*	0.012 (0.006)*	0.009 (0.006)
At least one parent has PSE below BA	0.024 (0.007)***	0.023 (0.007)***	0.020 (0.006)***	0.020 (0.006)***	0.016 (0.006)***
Both parents have a high school diploma	0.028 (0.011)***	0.027 (0.010)***	0.023 (0.009)**	0.024 (0.009)***	0.017 (0.008)**
One parent has a H.S. diploma	0.032 (0.011)***	0.030 (0.010)***	0.027 (0.009)***	0.024 (0.009)***	0.023 (0.010)**
Both parents have less than H.S.	0.078 (0.017)***	0.075 (0.016)***	0.071 (0.015)***	0.064 (0.015)***	0.065 (0.016)***
PISA scores and parents' aspirations– Reference PISA Quartile 4 and BA aspirations					
Below BA aspirations–PISA Quartile 1	0.137 (0.026)***	0.132 (0.026)***	0.102 (0.022)***	0.091 (0.020)***	0.073 (0.019)***
Below BA aspirations–PISA Quartile 2	0.073 (0.020)***	0.069 (0.019)***	0.051 (0.015)***	0.050 (0.016)***	0.035 (0.014)***
Below BA aspirations–PISA Quartile 3	0.045 (0.018)**	0.042 (0.017)**	0.031 (0.014)**	0.033 (0.014)**	0.031 (0.014)**
Below BA aspirations–PISA Quartile 4	0.022 (0.015)	0.020 (0.014)	0.019 (0.014)	0.017 (0.014)	0.006 (0.008)
BA and above aspirations–PISA Quartile 1	0.055 (0.016)***	0.051 (0.015)***	0.034 (0.011)***	0.039 (0.013)***	0.037 (0.013)***
BA and above aspirations–PISA Quartile 2	0.018 (0.007)**	0.017 (0.007)**	0.012 (0.006)*	0.012 (0.006)*	0.010 (0.006)*
BA and above aspirations–PISA Quartile 3	0.011 (0.007)*	0.011 (0.007)*	0.009 (0.006)	0.009 (0.006)	0.010 (0.006)*
Non-cognitive outcomes					
Child never 'just wants to get by'		-0.024 (0.008)***	-0.012 (0.007)*	-0.011 (0.007)	-0.010 (0.007)
Always does homework on time			-0.024 (0.007)***	-0.023 (0.007)***	-0.020 (0.007)***
Self-efficacy			-0.010 (0.004)**		
Self-esteem			-0.001 (0.003)		
Province dummies	Y	Y	Y	Y	Y
Family background	Y	Y	Y	Y	Y
Control for peers and behavioral outcomes	N	N	N	Y	Y
Control for school characteristics/local unemployment	N	N	N	N	Y
	N	N	N	N	Y
Sample size	7,755	7,755	7,661	7,574	6,245

continued, next page

Table 2.3: The effects of school characteristics, and non-cognitive and behavioral factors on dropping out for boys (cont'd) Marginal effects estimated in a Probit predicting dropping out at age 19. (Standard errors in parenthesis)

	1	2	3	4	5
Predicted probability of dropping out					
	0.037	0.035	0.030	0.029	0.023
Behavioral outcomes					
Youth reported a dependent child				0.166 (0.082)**	0.178 (0.090)**
Respondent smokes weekly age 15				0.052 (0.016)***	0.044 (0.015)***
Peer behavior **At at 15, all close friends:**					
Think completing high school is very important?				-0.006 (0.007)	-0.009 (0.006)
Skip classes once a week or more				-0.005 (0.016)	-0.005 (0.013)
Have dropped out of high school				0.046 (0.059)	0.041 (0.057)
Are planning education after high school				-0.017 (0.006)***	-0.016 (0.006)***
Have a reputation for causing trouble				0.004 (0.016)	0.005 (0.016)
Smoke cigarettes				0.005 (0.014)	0.001 (0.013)
Think it's okay to work hard at school				0.022 (0.013)*	0.019 (0.012)
School characteristics					
Low educational resources					0.016 (0.013)
Student to computer ratio					-0.001 (0.001)
Student to teacher ratio*10					0.0011 (0.0004)***
Local labor market					
Youth unemployment rate					-0.001 (0.001)
Province dummies	Y	Y	Y	Y	Y
Family background	Y	Y	Y	Y	Y
Control for *getby*	N	Y	Y	Y	Y
Control for *hmwrk*	N	N	Y	Y	Y
Control for other non-cognitive outcomes	N	N	Y	N	N
Sample size	7,755	7,755	7,661	7,574	6,245

Source: Youth in Transition Survey, Cycle 3 (Cohort A)
Estimates are weighted to account for non-response to the parents' survey and longitudinal attrition.
Standard errors clustered by high school.
*** indicates result is statistically significant at .01 level, ** at .05 level, * at .10 level
Marginal effects are evaluated for a youth whose parents both have a high school diploma at the mean of all the other variables.

Table 2.4: The effects of school characteristics, and non-cognitive and behavioral factors on dropping out for girls. Marginal effects estimated in a Probit predicting dropping out at age 19. (Standard errors in parenthesis)

	1	2	3	4	5
Predicted probability of dropping out					
	0.014	0.000	0.000	0.010	–
Log family income	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	–
*For a family of four, marginal effect of a difference in income from $50,000 and $18,000					
Parents' highest educational attainment –Reference both parents have a BA or higher					
One parent has BA	0.003 (0.006)	0.002 (0.006)	0.001 (0.006)	0.001 (0.006)	–
At least one parent has PSE below BA	0.012 (0.006)*	0.010 (0.006)*	0.008 (0.006)	0.006 (0.006)	–
Both parents have a high school diploma	0.009 (0.007)	0.008 (0.007)	0.004 (0.007)	0.005 (0.007)	–
One parent has a H.S. diploma	0.027 (0.009)***	0.025 (0.009)***	0.021 (0.009)**	0.016 (0.007)**	–
Both parents have less than H.S.	0.030 (0.010)***	0.028 (0.010)***	0.021 (0.009)**	0.019 (0.008)**	–
PISA scores and parents' aspirations– Reference PISA Quartile 4 and BA aspirations					
Below BA aspirations–PISA Quartile 1	0.065 (0.019)***	0.057 (0.018)***	0.043 (0.015)***	0.035 (0.013)***	–
Below BA aspirations–PISA Quartile 2	0.027 (0.012)**	0.024 (0.011)**	0.017 (0.008)**	0.017 (0.008)**	–
Below BA aspirations–PISA Quartile 3	0.023 (0.011)**	0.020 (0.009)**	0.015 (0.008)*	0.014 (0.008)*	–
Below BA aspirations–PISA Quartile 4	0.007 (0.007)	0.007 (0.006)	0.005 (0.005)	0.004 (0.005)	–
BA and above aspirations–PISA Quartile 1	0.045 (0.016)***	0.039 (0.015)***	0.031 (0.013)**	0.028 (0.011)**	–
BA and above aspirations–PISA Quartile 2	0.020 (0.008)**	0.018 (0.008)**	0.012 (0.006)**	0.012 (0.006)**	–
BA and above aspirations–PISA Quartile 3	0.000 (0.002)	0.000 (0.002)	0.000 (0.002)	-0.001 (0.002)	–
Non-cognitive outcomes					
Child never 'just wants to get by'		-0.012 (0.004)***	-0.006 (0.003)*	-0.005 (0.003)	–
Always does homework on time			-0.010 (0.004)***	-0.010 (0.004)**	–
Self-efficacy			-0.002 (0.001)		–
Self-esteem			0.000 (0.001)		–
Province dummies	Y	Y	Y	Y	
Family background	Y	Y	Y	Y	
Control for peers and behavioral outcomes	N	N	N	Y	
Control for school characteristics/local unemployment	N	N	N	N	
Sample size	8,376	8,376	8,329	8,239	

continued, next page

Table 2.4: The effects of school characteristics, and non-cognitive and behavioral factors on dropping out for girls (cont'd) Marginal effects estimated in a Probit predicting dropping out at age 19. (Standard errors in parenthesis)

	1	2	3	4	5
Predicted probability of dropping out					
	0.014	0.000	0.000	0.010	–
Behavioral outcomes					
Youth reported a dependent child				0.038 (0.020)*	–
Respondent smokes weekly age 15				0.015 (0.006)**	–
Peer behavior **At at 15, all close friends:**					
Think completing high school is very important?				-0.005 (0.003)*	–
Skip classes once a week or more				0.002 (0.012)	–
Have dropped out of high school				-0.009 (0.004)**	–
Are planning education after high school				-0.001 (0.003)	–
Have a reputation for causing trouble				0.023 (0.028)	–
Smoke cigarettes				0.000 (0.004)	–
Think it's okay to work hard at school				0.001 (0.004)	–
School characteristics					
Low educational resources					–
Student to computer ratio					–
Student to teacher ratio*10					–
Province dummies	Y	Y	Y	Y	
Family background	Y	Y	Y	Y	
Control for *getby*	N	Y	Y	Y	
Control for *hmwrk*	N	N	Y	Y	
Control for other non-cognitive outcomes	N	N	Y	N	
Sample size	8,376	8,376	8,329	8,239	

Source: Youth in Transition Survey, Cycle 3 (Cohort A)
Estimates are weighted to account for non-response to the parents' survey and longitudinal attrition.
Standard errors clustered by high school.
*** indicates result is statistically significant at .01 level, ** at .05 level, * at .10 level
Marginal effects are evaluated for a youth whose parents both have a high school diploma at the mean of all the other variables.

Table 2.5: Estimating the probability of dropping out among boys in a factor model.
Marginal effects reported (Standard errors in parenthesis)

	1	2
Log family income	-0.002 (0.001)	-0.003 (0.001)**
*For a family of four, marginal effect of a difference in income from $50,000 and $15,000		
Parents' highest educational attainment –Reference both parents have a BA or higher		
One parent has BA	0.021 (0.011)*	0.022 (0.022)
At least one parent has PSE below BA	0.026 (0.009)***	0.009 (0.020)
Both parents have a high school diploma	0.030 (0.011)***	0.007 (0.020)
One parent has a H.S. diploma	0.035 (0.011)***	0.007 (0.021)
Both parents have less than H.S.	0.075 (0.014)***	0.033 (0.022)
Other background characteristics		
Aboriginal	0.067 (0.024)***	0.052 (0.019)***
Immigrant	-0.021 (0.009)**	-0.019 (0.008)**
Rural	0.067 (0.024)***	0.052 (0.019)***
Number of moves	0.006 (0.007)	0.002 (0.001)***
Age in months	0.003 (0.008)	0.000 (0.001)
Number of siblings	0.001 (0.008)	-0.003 (0.002)
Local youth unemployment rate	-0.006 (0.008)	-0.010 (0.005)**
*Marginal effect of a difference in unemployment rate from top and bottom quartiles		
Family structure–Reference two biological parent families		
Other two parent families	0.074 (0.018)***	0.059 (0.012)***
Lone parent family	0.017 (0.012)	0.007 (0.007)
Province dummies	Y	Y
Factor distributions vary with parents' education	N	Y
Sample size	7,755	7,755

Source: Youth in Transition Survey, Cycle 3 (Cohort A)
Estimates are weighted to account for non-response to the parents' survey and longitudinal attrition.
Standard errors clustered by high school.
*** indicates result is statistically significant at .01 level, ** at .05 level, * at .10 level
Marginal effects are evaluated for a youth whose parents both have a high school diploma at the mean of all the other variables.
Column 1 estimated from a model where the probability weights are independent across parental education.
Column 2 estimated from a model where the probability weights differ across parental education.

Table 2.6: Predicted probability of dropping out conditional on ability and parental valuation. Factor distributions do not vary with parental education (Standard errors in parenthesis)

	High Non-Cognitive Skills					
	High Parental Valuation Cognitive Ability			Low Parental Valuation Cognitive Ability		
	High	Mid	Low	High	Mid	Low
Both parents have a BA	0.0000	0.0006	0.0113	0.0003	0.0087	0.0765
	(0.0000)	(0.0007)	(0.0094)	(0.0003)	(0.0067)	(0.0418)
One parent has BA	0.0001	0.0029	0.0353	0.0014	0.0282	0.1692
	(0.0001)	(0.0015)	(0.0133)	(0.0007)	(0.0078)	(0.0382)
At least one parent has PSE below BA	0.0001	0.0035	0.0405	0.0018	0.0326	0.1857
	(0.0001)	(0.0016)	(0.0121)	(0.0008)	(0.0062)	(0.0296)
Both parents have a high school diploma	0.0001	0.0042	0.0461	0.0021	0.0373	0.2024
	(0.0001)	(0.0022)	(0.0168)	(0.0011)	(0.0099)	(0.0407)
One parent has a H.S. diploma	0.0001	0.0051	0.0526	0.0026	0.0428	0.2209
	(0.0001)	(0.0025)	(0.0175)	(0.0013)	(0.0094)	(0.0375)
Both parents have less than H.S.	0.0006	0.0144	0.1082	0.0080	0.0909	0.3500
	(0.0004)	(0.0059)	(0.0283)	(0.0035)	(0.0169)	(0.0456)
	Low Non-Cognitive Skills					
	High Parental Valuation Cognitive Ability			Low Parental Valuation Cognitive Ability		
	High	Mid	Low	High	Mid	Low
Both parents have a BA	0.0000	0.0017	0.0236	0.0008	0.0186	0.1287
	(0.0001)	(0.0019)	(0.0186)	(0.0009)	(0.0132)	(0.0592)
One parent has BA	0.0002	0.0069	0.0653	0.0036	0.0536	0.2543
	(0.0002)	(0.0042)	(0.0255)	(0.0021)	(0.0141)	(0.0425)
At least one parent has PSE below BA	0.0003	0.0082	0.0737	0.0044	0.0608	0.2750
	(0.0003)	(0.0045)	(0.0246)	(0.0024)	(0.0122)	(0.0289)
Both parents have a high school diploma	0.0003	0.0097	0.0826	0.0052	0.0685	0.2957
	(0.0003)	(0.0057)	(0.0303)	(0.0029)	(0.0165)	(0.0409)
One parent has a H.S. diploma	0.0004	0.0115	0.0928	0.0063	0.0774	0.3182
	(0.0004)	(0.0066)	(0.0328)	(0.0034)	(0.0176)	(0.0398)
Both parents have less than H.S.	0.0015	0.0294	0.1736	0.0173	0.1494	0.4645
	(0.0014)	(0.0142)	(0.0478)	(0.0083)	(0.0270)	(0.0406)

Source: Youth in Transition Survey, Cycle 3 (Cohort A)
Estimates are weighted to account for non-response to the parents' survey and longitudinal attrition.
Standard errors clustered by high school.

Table 2.7: Predicted probability of dropping out conditional on ability and parental valuation. Factor distributions vary with parental education (Standard errors in parenthesis)

	High Non-Cognitive Skills					
	High Parental Valuation Cognitive Ability			Low Parental Valuation Cognitive Ability		
	High	Mid	Low	High	Mid	Low
Both parents have a BA	0.0000	0.0001	0.0023	0.0009	0.0133	0.0953
	(0.0000)	(0.0002)	(0.0030)	(0.0013)	(0.0126)	(0.0619)
One parent has BA	0.0000	0.0003	0.0058	0.0026	0.0282	0.1589
	(0.0000)	(0.0003)	(0.0040)	(0.0018)	(0.0122)	(0.0475)
At least one parent has PSE below BA	0.0000	0.0002	0.0036	0.0015	0.0191	0.1222
	(0.0000)	(0.0002)	(0.0026)	(0.0010)	(0.0070)	(0.0286)
Both parents have a high school diploma	0.0000	0.0001	0.0032	0.0014	0.0175	0.1152
	(0.0000)	(0.0002)	(0.0027)	(0.0010)	(0.0075)	(0.0317)
One parent has a H.S. diploma	0.0000	0.0001	0.0033	0.0014	0.0180	0.1174
	(0.0000)	(0.0002)	(0.0027)	(0.0010)	(0.0067)	(0.0271)
Both parents have less than H.S.	0.0000	0.0005	0.0083	0.0038	0.0377	0.1926
	(0.0000)	(0.0005)	(0.0061)	(0.0025)	(0.0126)	(0.0359)
	Low Non-Cognitive Skills					
	High Parental Valuation Cognitive Ability			Low Parental Valuation Cognitive Ability		
	High	Mid	Low	High	Mid	Low
Both parents have a BA	0.0000	0.0006	0.0103	0.0048	0.0447	0.2149
	(0.0000)	(0.0010)	(0.0115)	(0.0052)	(0.0326)	(0.0999)
One parent has BA	0.0001	0.0018	0.0224	0.0113	0.0824	0.3158
	(0.0001)	(0.0015)	(0.0141)	(0.0057)	(0.0224)	(0.0555)
At least one parent has PSE below BA	0.0000	0.0010	0.0149	0.0072	0.0601	0.2596
	(0.0001)	(0.0010)	(0.0102)	(0.0035)	(0.0117)	(0.0269)
Both parents have a high school diploma	0.0000	0.0009	0.0136	0.0066	0.0560	0.2483
	(0.0001)	(0.0010)	(0.0104)	(0.0036)	(0.0150)	(0.0373)
One parent has a H.S. diploma	0.0000	0.0010	0.0141	0.0068	0.0573	0.2519
	(0.0001)	(0.0010)	(0.0108)	(0.0037)	(0.0143)	(0.0333)
Both parents have less than H.S.	0.0001	0.0027	0.0303	0.0158	0.1041	0.3635
	(0.0002)	(0.0026)	(0.0205)	(0.0079)	(0.0228)	(0.0362)

Source: Youth in Transition Survey, Cycle 3 (Cohort A)
Estimates are weighted to account for non-response to the parents' survey and longitudinal attrition.
Standard errors clustered by high school.

Table 2.8: Distribution of factors, extended factor model (Standard errors in parenthesis)

	Cognitive Ability			Non-Cognitive Ability		Parental Valuation	
	High	Mid	Low	High	Low	High	Low
Both parents have a BA	0.520	0.418	0.062	0.332	0.668	0.848	0.152
	(0.041)	(0.040)	(0.016)	(0.033)	(0.033)	(0.045)	(0.045)
One parent has BA	0.328	0.492	0.180	0.364	0.636	0.730	0.270
	(0.030)	(0.030)	(0.021)	(0.030)	(0.030)	(0.043)	(0.043)
At least one parent has PSE below BA	0.165	0.508	0.328	0.380	0.620	0.544	0.456
	(0.019)	(0.021)	(0.022)	(0.026)	(0.026)	(0.036)	(0.036)
Both parents have a high school diploma	0.153	0.496	0.352	0.474	0.526	0.433	0.567
	(0.027)	(0.034)	(0.030)	(0.046)	(0.046)	(0.049)	(0.049)
One parent has a H.S. diploma	0.081	0.454	0.465	0.534	0.466	0.379	0.621
	(0.019)	(0.031)	(0.031)	(0.040)	(0.040)	(0.042)	(0.042)
Both parents have less than H.S.	0.078	0.341	0.581	0.543	0.457	0.300	0.700
	(0.016)	(0.030)	(0.031)	(0.045)	(0.045)	(0.044)	(0.044)

Source: Youth in Transition Survey, Cycle 3 (Cohort A)
Estimates are weighted to account for non-response to the parents' survey and longitudinal attrition.
Standard errors clustered by high school.

Table 2.9: Joint distribution of factors, extended factor model

	High Non-Cognitive Skills					
	High Parental Valuation Cognitive Ability			Low Parental Valuation Cognitive Ability		
	High	Mid	Low	High	Mid	Low
Both parents have a BA	0.146	0.118	0.017	0.026	0.021	0.003
One parent has BA	0.087	0.131	0.048	0.032	0.048	0.018
At least one parent has PSE below BA	0.034	0.105	0.068	0.029	0.088	0.057
Both parents have a high school diploma	0.031	0.102	0.072	0.041	0.133	0.095
One parent has a H.S. diploma	0.016	0.092	0.094	0.027	0.150	0.154
Both parents have less than H.S.	0.013	0.055	0.094	0.030	0.129	0.221
	Low Non-Cognitive Skills					
	High Parental Valuation Cognitive Ability			Low Parental Valuation Cognitive Ability		
	High	Mid	Low	High	Mid	Low
Both parents have a BA	0.295	0.237	0.035	0.053	0.043	0.006
One parent has BA	0.152	0.229	0.084	0.056	0.084	0.031
At least one parent has PSE below BA	0.056	0.171	0.110	0.047	0.143	0.093
Both parents have a high school diploma	0.035	0.113	0.080	0.046	0.148	0.105
One parent has a H.S. diploma	0.014	0.080	0.082	0.023	0.131	0.135
Both parents have less than H.S.	0.011	0.047	0.080	0.025	0.109	0.186

Source: Youth in Transition Survey, Cycle 3 (Cohort A)

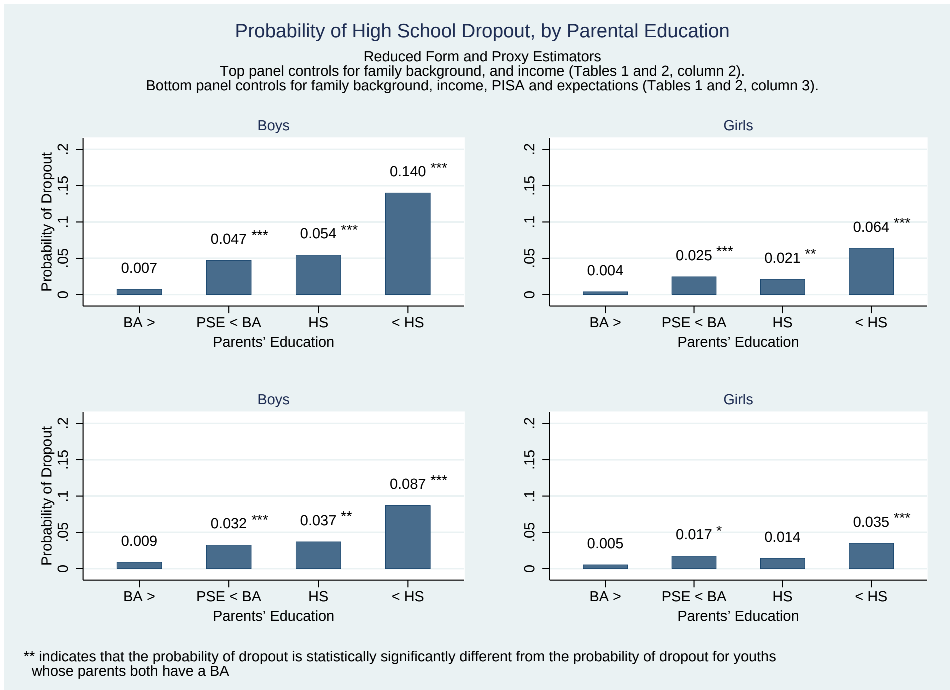

Figure 2.1: Probability of dropping out by parental education.

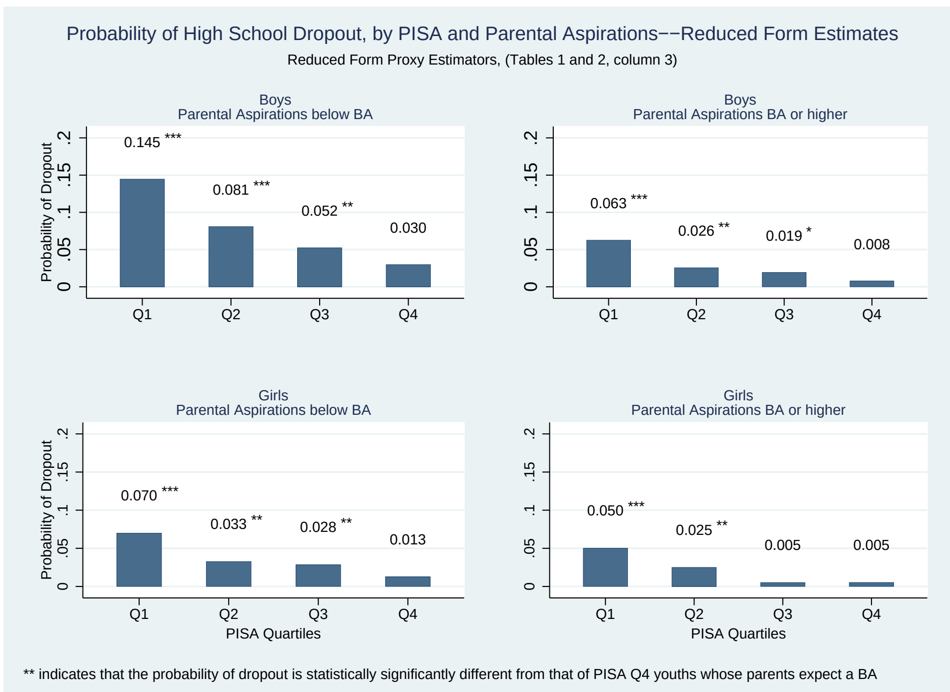

Figure 2.2: Probability of dropping out by PISA and parental aspirations – Reduced Form.

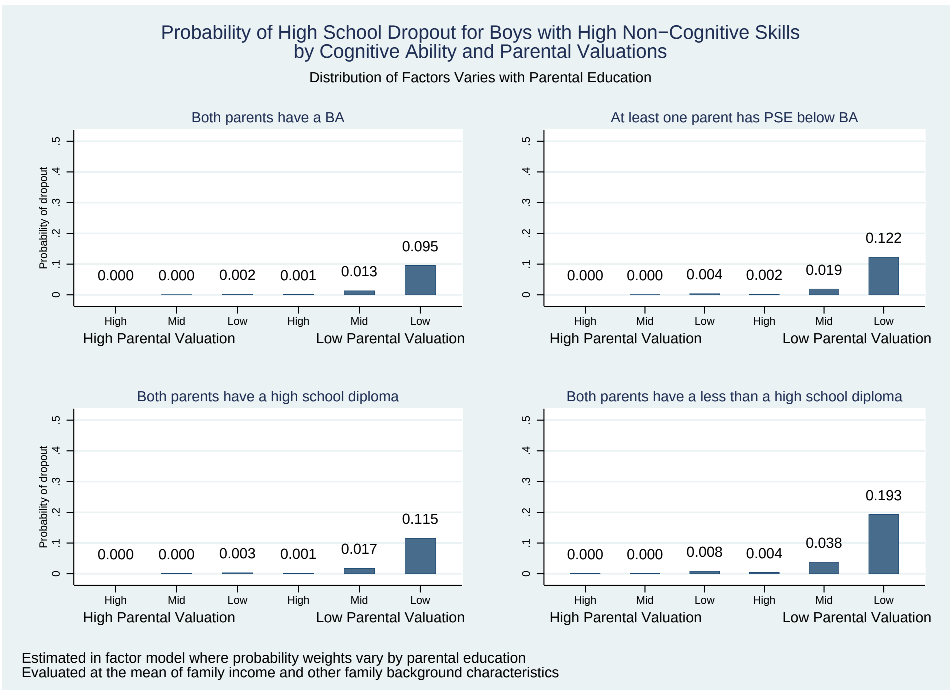

Figure 2.3: Probability of dropping out by cognitive ability and parental valuations for boys with high non-cognitive skills – extended factor model.

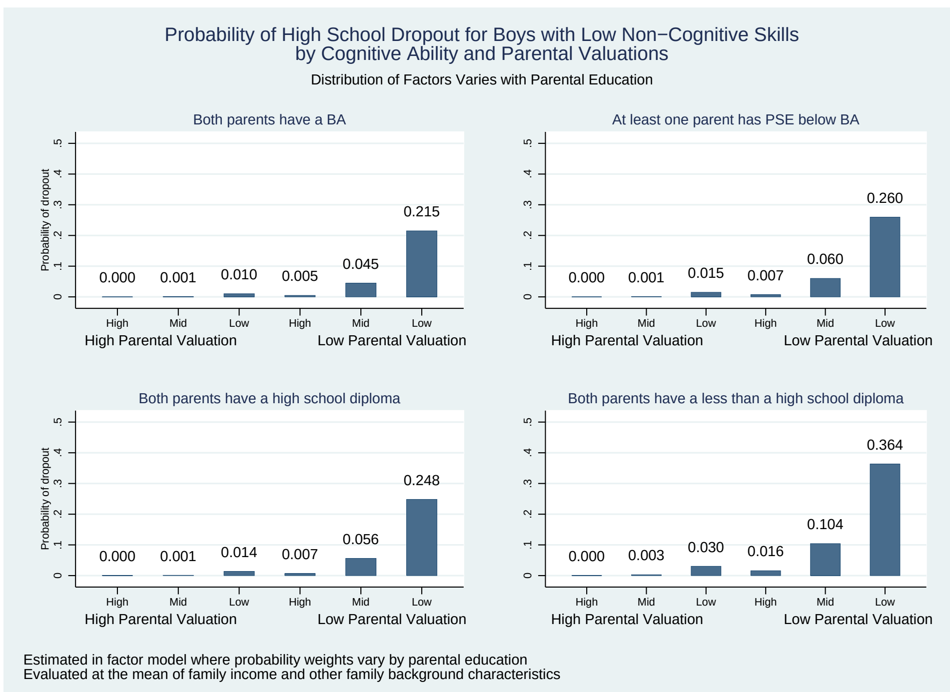

Figure 2.4: Probability of dropping out by cognitive ability and parental valuations for boys with low non-cognitive skills – extended factor model.

2.7 Bibliography

K. Alexander, B. Eckland, and L. Griffin. The wisconsin model of socioeconomic achievement: A replication. *The American Journal of Sociology*, 81(2):324–342, 1975.

O. Attanasio and K. Kaufmann. Educational choices, subjective expectations, and credit constraints. NBER Working Papers 15087, National Bureau of Economic Research, Inc, July 2009.

J. Behrman and M. Rosenzweig. Does increasing women's schooling raise the schooling of the next generation? *American Economic Review*, 92(1):323–334, 2002.

P. Belley, M. Frenette, and L. Lochner. Educational attainment by parental income: A canada - u.s. comparison. mimeo, Dept of Economics, University of Western Ontario, 2008.

L. Borghans, A. L. Duckworth, J. J. Heckman, and B. ter Weel. The economics and psychology of personality traits. *Journal of Human Resources*, 43(4), 2008.

P. Carneiro and J. Heckman. Human Capital Policy. *NBER Working Paper*, 9495, 2003.

P. Carneiro, K. T. Hansen, and J. J. Heckman. Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. *International Economic Review*, 44(2): 361–422, 2003.

P. Carneiro, C. Meghir, and M. Parey. Maternal education, home environments and the development of children and adolescents. Wp15/07, Institute for Fiscal Studies, 2007.

G. Chamberlain. Education, income and ability revisited. *Journal of Econometrics*, 5(2):241–257, 1977.

M. Coelli. Parental job loss, income shocks and the education enrolment of youth. Department of Economics Working Paper 1060, The University of Melbourne, 2009.

F. Cunha and J. Heckman. The technology of skill formation. *American Economic Review*, 97(2):31–47, May 2007.

F. Cunha and J. J. Heckman. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, 43(4), 2008.

F. Cunha, J. Heckman, and S. Schennach. Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. mimeo, Dept of Economics, University of Chicago, 2006.

M. Davies and D. B. Kandel. Parental and peer influences on adolescents' educational plans: Some further evidence. *Review of Economic Studies*, 87(2): 363–387, 1981.

de Broucker P. Without a paddle: What to do about canada's young drop-outs. Report w/30, Canadian Policy Research Network, 2005.

R. Dworkin. What is equality? part 1: Equality of welfare. *Philosophy and Public Affairs*, 10:185–246, 1981a.

R. Dworkin. What is equality? part 2: Equality of resources. *Philosophy and Public Affairs*, 10:283–345, 1981b.

Z. Eckstein and K. I. Wolpin. Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica*, 67(6):1295–1340, November 1999.

G. Gallipoli, C. Meghir, and G. Violante. Equilibrium effects of education policies: a quantitative evaluation. UCL Working paper, 2009.

R. Haveman, B. Wolfe, and J. Spaulding. Childhood events and circumstances influencing high school completion. *Demography*, 28(1):133–157, 1991.

J. Heckman and B. Singer. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52(2): 271–320, March 1984.

J. J. Heckman and L. Lochner. Rethinking education and training policy: Understanding the sources of skill formation in a modern economy. In S. Danziger and J. Waldfogel, editors, *Securing the Future: Investing in Children from Birth to College*, pages 47–83. New York: Russell Sage Foundation, 2000.

J. J. Heckman, J. Stixrud, and S. Urzua. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482, July 2006.

S. Junor and A. Usher. The price of knowledge 2004: Access and student finance in canada. Canadian Millennium Scholarship Foundation, 2003.

P. Pintrich and E. D. Groot. Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1): 33–40, 1990.

M. Rosenberg. *Society and the Adolescent Self-Image*. Princeton, New Jersey: Princeton University Press, 1965.

W. Sewell, O. Haller, and A. Portes. The educational and early occupational attainment process. *American Sociology Review*, 34:82–91, 1969.

P. E. Todd and K. I. Wolpin. The production of cognitive achievement in children: Home, school and racial test score gaps. University of Pennsylvania, Working Paper, 2006.

Chapter 3

Can Neighbourhoods Change the Decisions of Youth on the Margins of University Participation?[1]

3.1 Introduction

What determines university participation is an issue of great interest to economists and policy makers, and for good reason. Educational attainment is closely linked with many benefits for individuals and society, including higher earnings and income, better health, lower exposure to crime, and greater civic participation. A large literature studying the determinants of university participation has developed over the last two decades. Central to this literature is a vigourous debate on the specific role that family background plays. Some researchers argue that credit constraints prevent low-income children from attending university (Belley and Lochner 2007, Keane and Wolpin 2001), while others argue that credit con-

[1] A version of this chapter will be submitted for publication. K. Foley. Can Neighbourhoods Change the Decisions of Youth on the Margins of University Participation?

straints are minimal and the most important barrier associated with low income stems from growing up in a developmentally impoverished environment. As Carneiro and Heckman (2002) suggest, there is no market for families and the constraint arises from "the inability of the child to buy the parental environment and genes that form the cognitive and noncognitive abilities required for success in school," (Carneiro and Heckman 2002; 706).

The family and home environment is not, however, the only arena in which young people develop. Schools and neighbourhoods provide extra-familial contexts where interactions with peers, friends and other adults might mitigate or reinforce family impacts. This paper is interested in whether the local neighbourhood in which a youth resides can affect her chances of participating in university. Neighbourhood influences are measured by the share of adults living in close proximity who have a Bachelors degree. The neighbourhood BA share proxies the availability of adult role models and peers who are oriented toward university.

An important part of the empirical literature estimating neighbourhood effects has found modest or no impacts on a range of outcomes (Oreopoulos 2003, Kling et al. 2005). This paper contributes to that literature by demonstrating that neighbourhoods can have substantial effects for individuals who are near the margins of economic decisions. I take advantage of a rich Canadian data set, the Youth in Transition Survey (YITS), to estimate how neighbourhoods change the marginal probability of university attendance at different points on the distributions of skills and socio-economic background.

The YITS is a longitudinal survey of youth from across Canada, and includes a standardized reading test and a parental survey which collected information on family income, education and the level of education parents hope their children will achieve. A key feature of the data used in this paper is detailed information about where youth attended high school and where they lived at age 15.

Estimating the relationship between neighbourhoods and individual outcomes is particularly difficult because families with similar unobserved characteristics tend to cluster together, making it difficult to disentangle the effects of family from the neighbourhood impacts. To identify the neighbourhood effect, I assume that, conditional on school choice, family background and a measure of reading skills, selection into neighbourhoods is uncorrelated with unobservable factors affecting children's university participation. The assumption is grounded in evidence from the literature which suggests that school quality has value for families. If parents focus on schools when choosing their neighbourhood, one would expect considerable sorting across schools. I assume variation in neighbourhood BA share, conditional on high school, family background and a measure of reading skills, is exogenous because sorting *within* school catchments is constrained by the thinness of markets for houses in specific neighbourhoods. A neighbourhood in this paper is a geographic area of roughly one to four blocks in size. The ability of a family to locate in a specific neighbourhood is constrained by the availability of a suitable house within the family's search period. Neighbourhood mobility rates, estimated from data in the Canadian Census, suggest that to be guaranteed a home in a specific neighbourhood, families would need to wait 8 years.

I find that neighbourhoods are strongly related to university participation. Compared to a youth who lives in a community with no university-educated adults, a youth living in a community where all adults hold a Bachelors degree is roughly 24 percentage points more likely to attend university. At the mean, this effect would increase the probability of attending university by 72 per cent for boys and 49 per cent for girls.

I also find that marginal effects on university participation differ by family background and reading skills, measured by scores on a standardized test. For youth with a university educated parent, neighbourhoods have no impact on participation independent of reading skills. In contrast, for those with less educated parents, neighbourhood effects differ across the distribution of reading skills. Neighbourhoods have a statistically significant but economically tiny effect on participation among low-skilled youth from less educated, low-income families. The marginal effect is largest for disadvantaged youth in the highest skills quartile. This suggests that highly skilled youth are most affected by influences in their neighbourhood when they come from disadvantaged backgrounds.

The rest of the paper proceeds by first describing reasons from the theoretical literature why neighbourhoods might affect individual behaviour. The empirical evidence of neighbourhood effects is then reviewed. Section 2 discusses the empirical issues associated with estimating neighbourhood effects and outlines the identification strategy. The data are described in Section 3. The results section follows and includes a discussion of the main neighbourhood effects and presents alternative analysis and samples to demonstrate the robustness of the results. Finally, the paper reports the estimated marginal effects evaluated at different points along measures of skills and socio-economic background.

3.1.1 The Theoretical Link between Neighbourhoods and Individual Behaviour

Residential neighbourhoods are potential forums for social interaction. In the last two decades, economists have added to the growing psychology literature that explores how individuals are affected by group interaction, behaviour and beliefs.[2] Two broad effects are identified in this literature. The first is peer or conformity effects, which describes the tendency for individuals to behave like group members because they experience some positive utility from group membership. Akerlof and Kranton (2000), for example, develop a model in which individuals have a sense of identity that is associated with membership in a particular group. Engaging in behaviour which undermines that identity reduces utility, while behaviour which supports identity increases utility. In Bernheim (1994), individuals achieve status, which is desirable, through peers' subjective assessment of their predispositions rather than their actions. Society penalizes individuals for non-conformity which gives rise to homogenous behaviours even when underlying preferences are heterogeneous.

In the context of university participation, one can think of neighbourhoods as exhibiting different cultural propensities for education. This propensity might manifest itself in language, for example. Among local youth the use of formal diction and correct grammar might be ridiculed. A high school student who was preparing

[2]Durlauf (2004) provides a survey of the theoretical literature on neighbourhood and group effects.

to attend university would find it difficult to simultaneously fit in with peers and develop skills necessary for success in higher education.[3]

The literature on evolutionary dynamics in cultural transmission provides some clues as to how some groups might develop particular preferences for or against education. Bisin and Verdier (2001) show that heterogenous group preferences can exist when parental and extra-familial socialization are substitutes. The model assumes that parents socialize their children according to their own preferences because the altruistic assessment of children's welfare is undertaken from the parents' perspective. The result, that heterogenous group preferences can exist, contrasts with simple evolutionary selection models without interaction among individuals where equilibria are homogenous.

The second broad category of group effects can be described as role model effects, which arise from incomplete information. When information is not fully or easily available, individuals may either gain information from observing those around them or may simply mimic behaviour for lack of information.

This effect may be particularly important with respect to university participation. Survey evidence in Canada suggests that low income youth and their parents over-estimate the costs of education and under-estimate the benefits (Junor and Usher 2004). Using U.S. data from the NLSY, Ludwig (1999) similarly found that youth from low-income urban areas have inaccurate information about the educational requirements of their occupational goals.

While the literature provides convincing theoretical reasons why neighbourhoods should matter, as Oreopoulos (2008) points out, one should also consider why neighbourhoods might not matter. He argues that neighbourhoods may not be the most important venue for social interactions. Individuals might ignore their neighbours, drawing instead from work or other social networks.

While this is certainly true of adults, children have a geographic scope which is relatively limited. Some children may be exposed to 'play groups' or other activities outside their local area; however, for children from lower socio-economic

[3] This argument is similar to a concept described as 'acting white' in the U.S. literature (Fryer and Torelli 2005).

backgrounds this would be less likely. As children begin to engage in autonomous and independent activities, they are most likely to be restricted to their local areas. If identity is developed at an early age and persists then relatively small geographic localities may have enduring impacts on individuals. This is relevant to keep in mind since the geographic space which defines a neighbourhood is fairly small in this paper.

For youth, beyond family and neighbourhood, schools are probably the most important social arena. The Coleman Report 1966, a seminal work in the early sociological literature on low-income neighbourhoods in the U.S., suggested that schools were less important relative to the effects of inner-city "ghettos". More recent work by Card and Krueger (1996) counters this view with evidence that school quality affects labour market outcomes in adulthood. Duncan et al. (2001) estimate the relative strength of associations between siblings, neighbours and schoolmates in the same grade. Although much smaller than the within-family correlations, correlations among neighbours are similar or larger than those among schoolmates.

The model I estimate in this paper does not require that local neighbourhoods are the only social forum which affects behaviour. Instead, what matters is that the social influences in a youth's neighbourhood are incrementally important, holding constant school and family effects.

Unfortunately, it is not possible with the data available to identify the specific mechanism that generates the relationship between neighbourhoods and university participation. The possible mechanisms include those outlined in the literature such as peer or conformity effects and role model or informational effects. For example, youth might be targeted for bullying by their neighbours for showing an interest in school. Alternatively, children might spend time in the homes of their neighbours learning about other professions from the parents of their friends who live nearby.

3.1.2 Empirical Evidence of Neighbourhood Effects

Although theory clearly predicts an important role for neighbourhoods, it has proven extremely difficult to empirically establish a casual link between neighbourhoods and economic outcomes. Over the past three decades, many researchers have attempted to estimate the impact of neighbourhoods. Oreopoulos (2008) reviews the Canadian literature, while (Vigdor 2006) comprehensively surveys the U.S. literature. [4]

The key difficulty in establishing causality lies in distinguishing neighbourhood effects from the effects of family background. For various reasons, certain types of families tend to cluster together, creating a correlation between neighbourhood and family characteristics.

Many empirical studies have relied on observational data and multivariate regressions to estimate neighbourhood effects. The research using this type of approach has generally found substantial neighbourhood impacts on a range of outcomes including income, education (Borjas 1995, Gibbons 2001), and childhood development and health (Brooks-Gunn et al. 1993, Kohen et al. 2002, Curtis et al. 2004, Tremblay et al. 2001), and social assistance receipt (Corcoran et al. 1992).

In contrast, the branch of literature using experimental and quasi-experimental methods report zero or modest neighbourhood effects. Two prominent housing policy interventions that researchers have used to test for neighbourhood effects include the Moving to Opportunity (MTO) project (Orr et al. 2003) and the Metro Toronto Housing Corporation (Oreopoulos 2003).

The MTO project, which was evaluated using a random assignment design, offered low-income residents of public housing a voucher and support to move to higher income neighbourhoods. Although on many dimensions of quality, MTO led participants to live in better neighbourhoods, there were no impacts on economic outcomes for adults and the children's schooling outcomes were modest (Orr et al. 2003). Living in better neighbourhoods did not increase earnings and employment and did not reduce social assistance receipt among adults. On average, children attended schools that were marginally better in terms of scores on state exams. However, many children did not change schools because their parents' moved or remained in the same school district. Educational achievement and

[4] Related to this literature is the peer effects literature, which considers the impact of school classmates' behaviour on children's outcomes (see Duncan et al. 2005, Hoxby 2000; for example). This literature contends with many of the same issues related to estimating neighbourhood effects and often estimates impacts that combine neighbourhood and peer influences.

performance were not improved.

Some important gender differences were found among adolescents. Significant improvements were found for girls in terms of mental health outcomes and some risk taking behaviours (Kling et al. 2005). In contrast, boys were actually more likely to report marijuana use in the previous month and were more likely to have been arrested for property crime.

In the second example, Oreopoulos (2003) exploited variation in neighbourhood quality arising from the Metro Toronto Housing Corporation's allocation policy. Candidates for public housing in Toronto were placed on a lengthy wait list. When housing that suited their family size became available, applicants were assigned to housing projects irrespective of their preferences. The projects were located in neighbourhoods varying in quality.

To estimate the effect of neighbourhood quality, Oreopoulos used administrative data from the Intergenerational Income Database (IID). The IID provides data on tax returns filed between 1978 and 1999 for children born between 1963 and 1970. Oreopoulos (2003) defined quality alternatively by public housing density, public housing size (number of units), percentage of census tract with income below the Low Income Cut-Off, and by housing type (townhouse or apartment). Along these dimensions of quality, Oreopoulos (2003) found no significant differences in any economic outcomes.

Although policy interventions can have the advantage that they generate exogenous variation in neighbourhood, the neighbourhood impacts are possibly confounded by the nature of the intervention. In other words, the research tests the effect of neighbourhood in combination with an intervention. In MTO, the intervention offered not only a different type of neighbourhood but also a household move and entry into the private housing market. The natural experiment in the Oreopoulos (2003) study similarly combined different neighbourhood qualities with moving from private to public housing.

Ethnographic research suggests that policy induced residential moves may not actually change an individual's social group. While chronicling the demise of Chicago's Robert Taylor public housing project, Venkatesh (2000a) described how some of the families who moved from the project to better neighbourhoods were followed by abusive partners or gang-involved family members. Indeed, a transition study estimated that roughly two-thirds of household heads reported persistent abuse or harassment from previously co-habitant partners (Venkatesh 2000b).

In trying to understand gender differences in the MTO results, Clampet-Lundquist et al. (2006) used data from in-depth interviews to explore coping strategies employed by experimental and control group members. They found that many boys in the experimental group returned to their low-income neighbourhoods, while control group boys actively employed strategies to avoid trouble.

Another set of studies employs instruments and other types of plausibly exogenous neighbourhood variation. Aaronson (1998) for example, compared sibling outcomes among families who had moved. The residential move provided variation in neighbourhood while family characteristics remained constant. Aaronson (1998) found large neighbourhood impacts on the probability of high school dropout. He estimated that a 10 percentage point increase in the local poverty rate reduces the chance of high school completion by 7 percentage points.

Card and Rothstein (2007) tested whether racial segregation in neighbourhoods and schools can explain the black-white gap in SAT scores. They aggregated across metropolitan areas, which eliminated the correlation between neighbours' characteristics and their preferences for integrated neighbourhoods. Any metropolitan level effects were eliminated by differencing across racial groups. The remaining variation in racial segregation across cities was then related to the black-white differential in test scores. Their results suggest that changing from a fully segregated to a fully integrated city would reduce the test score gap by 25 per cent. The estimated effects are larger for neighbourhood segregation compared to school segregation and are associated with variation in income rather than race.

With data from the French Labour Force Survey, Goux and Maurin (2007) examine whether the proportion of youth living in close proximity who have repeated a grade at age 15 will affect the chances that an individual repeats a grade by age 16. The authors use the month of birth to instrument for neighbours' educational advancement. They argue that because it determines when children begin school the distribution of birth months is related to early childhood achievement but is unlikely to be related to later school outcomes. They find that a one standard deviation increase in the proportion of neighbours who have been held back a grade will increase the chance that a youth is held back by 10 to 15 percentage points.

This paper contributes to this literature by studying how neighbourhoods affect individuals on the margins of an important economic decision. In contrast to many studies which focus on disadvantaged samples, I use data from across the socio-economic distribution. This allows me to explore what characterizes an individual for whom neighbourhood influences might make the determining impact on university participation. Understanding neighbourhood effects for the marginal individual is important for two reasons. First, this information can help policy makers target policy to improve its cost-effectiveness. Second, it demonstrates empirically that the estimated impact of neighbourhoods on a discrete outcome can be very small at some points on the socio-economic distribution while large and economically substantial at other points.

3.2 Estimation and Identification

One of the key challenges in identifying the impact of neighbourhoods stems from the fact that families are not randomly distributed across their neighbourhoods. Similar families tend to cluster together generating a correlation between unobserved family characteristics and neighbourhood characteristics. To disentangle the effects of neighbourhoods from the impact of the family, I compare the probability of university attendance among similar individuals attending the same high school.

This approach is very similar to one employed by Bayer, Ross, and Topa (2005), who are interested in whether local neighbourhoods provide effective job referral networks. Using data from the U.S. Census, Bayer et al. (2005) compare the probability that an individual works with a neighbour living on the same city block in metropolitan Boston, to the probability of working with a neighbour from nearby blocks. They argue that while individuals do choose their residential area, defined as a block group consisting of roughly 10 blocks, sorting within block groups is minimal because the market for a specific block is too thin. They find that people living on the same U.S. Census block are 33 per cent more likely to work together when compared to those within the block group.

In a similar fashion, I use variation in neighbourhood BA share at the scale of one to four blocks and argue that the geographic scale of this variation is smaller than the scale at which, in practice, families can choose where they live. In other words, while families may have a preference for a specific block, the chances are small that a house which suits their specific needs and tastes is actually for sale during the family's search period. Estimates from the Canadian Census suggest that if a family wanted to live on a specific block, they would have to wait approximately 8 years to be guaranteed at least one house came up for sale.[5]

In practice, to ensure that they can find a house, the typical family would search for houses and would be willing to live within geographic areas that are much larger than a collection of one to four blocks. Because a large part of why families choose where they live is the quality of the schools their children will attend, a school catchment area is a reasonable way to think about a family's search area.

In the U.S., several studies confirm that families care about school quality when they choose their homes. Specifically, the literature suggests that school quality is capitalized into housing prices. Bayer, Ferreira, and McMillan (2007) found that housing prices in San Francisco increased across school catchment boundaries, even though measures of housing quality were continuous across the boundary. They estimated that households will pay less than one percent more for a 5 percent increase in school quality. Estimates comparing relocation choices of childless families and those with children also showed that schools have substantial value Barrow (2002).

In Canada, while very little research has addressed the role of schools in fami-

[5] I use Enumeration Areas (EA) as the geographic boundaries of a neighbourhood. EAs are the smallest geographic classification used by Statistics Canada and are roughly one to four blocks in size. At the median EA in the sample, roughly 12 per cent of the households had moved between 1999 and 2000. The figure of 8 years comes from assuming that mobility rates are uniform across EAs.

lies' location decisions, there is some evidence that schools are similarly important for Canadian families. One example is work by (Ries and Somerville 2004) who exploited a change in school boundaries which meant a group of houses in Vancouver formerly assigned to a low quality school catchment were reassigned to a substantially higher quality school. Their estimates suggest that a 4 per cent change in high school quality is associated with an increase in the median house price of $14,000.

What this suggests in terms of identifying neighborhood effects is that children from families with similar income and who attend the same school have parents with similar preferences for school quality. In essence, identification comes from the assumption that conditional on family income, parental education, children's reading skills and attending the same high school, the unobserved determinants of university participation are uncorrelated with neighbourhood characteristics.

Although it is not possible to definitely point to the source, as it would be the case in with a policy intervention, it is worth suggesting the possible mechanisms that could generate exogenous variation in these circumstances. First, there may be heterogeneous preferences for housing characteristics that are unlikely to be correlated with university participation. For example, preferences for ranch verses Victorian style homes. Heterogenous demand does not necessarily lead to higher prices. Consequently, families with similar wealth might be distributed across neighbourhoods according to those preferences.

Another possible source of exogenous variation conditional on schools and income is the process that generates houses on the market. Houses become available in a specific neighbourhood for many reasons that are uncorrelated with the neighbourhood's characteristics. These reasons would include any idiosyncratic family shocks, such as divorce or promotion which requires relocation. Naturally, houses are also available on the market for endogenous reasons, such as declining labour markets or rising crime rates. Such trends in neighbourhood quality might cause sorting across school catchment areas but are unlikely to be so localized as to affect only a few blocks.

Finally, variation will exist within a school catchment area because in any given city or region, the menu of school and housing quality choices will not be continuous. If there were as many schools to choose from as there were family-preference types then families would perfectly sort by school and there would not be any neighbourhood variation within catchment areas. It is more reasonable to assume that there are many more preference types than there are choices over school and housing quality. Because of the range of schools catchments within a given city one might, for example, observe variation in neighbourhood BA share because a plumber's family lives beside a school teacher's family.

Because in essence, I identify a neighbourhood effect by invoking a conditional independence assumption it always possible that some unobserved factor is correlated with neighbourhood characteristics. In a later section, I will show that the pattern of results are inconsistent with correlation between the variation in neighbourhood BA share and plausibly relevant unobserved factors, such as parents' ambitions for their children's educational achievement.

3.2.1 Empirical Model

In the next section, I write down the empirical model with the particular aim of describing more precisely what is identified in the estimation. I follow Manski (1993), Moffitt (2001) and Brock and Durlauf (2007), who have described and summarized the issues associated with estimating group interaction effects.

Individuals decide to attend university when they expect their lifetime utility with university to exceed their lifetime utility without university. The net utility for individual i, living in neighbourhood g, attending school s can be defined as:

$$\omega_{igs} = x_i\beta^* + x_s\psi^* + x_g\gamma^* + \bar{\omega}_g\theta^* + \mu_g^* + \mu_s^* + \varepsilon_{igs}^* \tag{3.1}$$

This definition suggests that net utility is a function of observable individual and family characteristics (x_i), observable (x_s) and unobservable (μ_s^*) school characteristics, a set of neighbourhood effects and an idiosyncratic error (ε_{igs}^*). The neighbourhood effects can be decomposed into three parts. The first is the effect of observable characteristics (x_g), which is often referred to as an exogenous or contextual effect. The second part is referred to as the endogenous effect, and is the impact of $\bar{\omega}_g$, which is defined as $\mathbb{E}\left[\omega_{igs}|x_g,\mu_g^*\right]$ or the individual's subjective expectation of group preferences. The final component is unobserved neighbourhood characteristics μ_g^*.

In a linear model, the fact that $\bar{\omega}_g$ is a linear function of the other regressors gives rise to what (Manski 1993) called the reflection problem, which prevents identification of γ^* and θ^* without additional structure. Brock and Durlauf (2007) prove that in non-linear models, such as discrete choices, under some assumptions and given sufficient variation, these parameters can be identified from the non-linearity. There are not enough observations in the data within each neighbourhood to estimate $\bar{\omega}_g$.[6] For that reason, I estimate a reduced form:

$$\omega_{igs} = x_i\beta + x_s\psi + x_g\gamma + \mu_g + \mu_s + \varepsilon_{igs} \tag{3.2}$$

The parameter of interest is now $\gamma = (\gamma^* + \beta^*\theta^*)/(1-\theta^*)$. If γ is non-zero then it means that either γ^* or $\beta^*\theta^*$ is non-zero. It is not possible to distinguish whether any estimated neighbourhood effect stems from the choices of peers or from contextual factors, such as adult role models. Because I compare youth who attend the same high school many potentially relevant contextual factors, such as teacher quality, will not be reflected in the estimate of γ.

While net utility determines university participation, I observe only the outcome in my data:

$$uni_{igs} = \begin{cases} 1 & \text{if } x_i\beta + x_s\psi + x_g\gamma + \mu_g + \mu_s + \varepsilon_{igs} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

The process that determines the school and neighbourhood in which I observe a particular individual is not random and is instead the outcome of her family's decision of where to live. Taking into account neighbourhood and school char-

[6] I use the fraction of adults in the neighbourhood with a BA to proxy x_g. Measures of neighbourhood characteristics come from the 2001 Canadian Census data. A reasonable measure of μ_g is not available in the Census data.

acteristics as well as their children's observable and unobservable characteristics, families choose neighbourhoods from within a set of affordable houses to maximize their utility:

$$U_{igs} = \mathbf{U}(x_i, x_g, x_s, \mu_g, \mu_s, \varepsilon_{igs}) + \upsilon_{igs} \tag{3.4}$$

Because families take into account their children's unobserved characteristics and because certain types of families might have similar preferences, the first-best neighbourhood choice or $\mathbb{P}[U_{igs} \geq U_{iht} \;\; \forall \, h \, t]$ is correlated with the error in (3.3). This means that $\mathbb{E}[\varepsilon_{igs} | x_i, x_g, \mu_g] \neq 0$.

A family is observed in a neighbourhood not just because that family prefers the particular neighbourhood but also because a house within their budget set is available when the family is searching. I assume that conditional on school selection the process which generates houses on the market within a particular budget set is uniform across neighbourhoods and independent of $\mu_g + \varepsilon_{igs}$. As discussed earlier in Section 2, this assumption is based on the idea that any endogenous factors that make houses available, such as trends in crime rates or labour market returns, are not localized in a specific neighbourhood. Under these conditions, $\mathbb{E}[\mu_g + \varepsilon_{igs} | x_i, x_s, \mu_s] = \mathbb{E}[\mu_g + \varepsilon_{igs} | x_i, x_g, x_s, \mu_s] = 0$.

To estimate the model, I assume that $\mu_g + \varepsilon_{igs}$ is normally distributed and estimate the probability of university attendance in a Probit,[7] with school fixed effects α_s which will absorb both $x_s \psi$ and μ_s:

$$\mathbb{P}[uni = 1 | x_i, x_g, s] = \Phi(x_i \beta + x_g \gamma + \alpha_s) \tag{3.5}$$

What is of particular interest in this paper, is not just whether γ is non-zero but also if γ varies across family types, reading skills and whether the neighbourhood effect is large enough to change the outcome of university participation. To test whether neighbourhood effects differ across background characteristics I also

[7]In general, fixed effects in a Probit do not consistently estimate the parameters. If however, one assumes that α_s is proportional to the mean of the independent variables then average partial effects are consistently estimated (Wooldridge 2002). Moreover, the estimates in linear probability models are very similar.

estimate (3.5) with interactions between x_i and x_g.

It is possible, and probable in light of the importance of family background, that even if γ is non-zero, for some youth, neighbourhoods will have no impact on the marginal probability of university participation. One of the contributions of this paper is to examine at which point along the distribution of net utility the marginal effect of neighbourhoods on university participation is largest. Clearly, net utility is not observed in data available to researchers or policy makers. I use instead observable measures of youths' skills, family background, and parental hopes for their children's education. These measures are likely to be correlated with net utility and could be used by policy makers to target programmes and services.

3.3 Data

This paper uses data from the Youth in Transition Survey (YITS). The YITS is a longitudinal survey of Canadian youth undertaken by Statistics Canada in partnership with Human Resources and Skills Development Canada. While two cohorts of youth were surveyed in the YITS project, this paper uses data from the younger cohort (YITS-A). The youth were born in 1984, were surveyed first in 2000, and again every two years. Four cycles of data were available for use in this paper. The YITS data is particularly rich because the first cycle for the younger cohort included an internationally comparable reading test, a parental survey and a school administrators survey.

The reading test was administered through the Programme for International Student Assessment (PISA). PISA was an effort, co-ordinated by the Organization for Economic Co-operation and Development (OECD), to generate internationally coherent measures of cognitive skills.

The parent survey was completed by the parent or guardian who identified him or herself as "most knowledgeable" about the child.[8] The responding parent provided data about their education, work, and income. Parents also answered questions about their aspirations for their children's education.

[8]The responding parent is not necessarily a birth or adoptive parent even when there is a birth or adoptive parent in the household.

The sample used in this paper is drawn from the 'reading cohort' and includes youth who were respondents to all four waves of the survey and whose parent completed the parental survey in the first cycle.[9] Youth who board at school or commute farther than 45 minutes by car or public transport are excluded. These youth are dropped from the sample because it is less likely that their residential location is linked to their school attendance.

The base sample is 13,611, which represents roughly 45 per cent of the original sample. Most of the sample is lost due to longitudinal attrition. The overall survey response rate is 54.7 per cent. Weights provided by Statistics Canada are used to account for attrition. Other survey respondents are excluded from the analysis because of missing data in the co-variates or inability to match neighbourhood data to the YITS data.

3.3.1 Dependent Variable

The outcome of interest is university participation and is measured when the youth are age 21. Survey respondents who reported ever attending a university prior to December 2005 are defined as university participants. The university participation rate in the sample is roughly 40 per cent. Girls are much more likely to have attended university. The participation rate is 48 per cent among girls and 33 per cent among boys.

Because some individuals may attend university later in their life, this measure will underestimate participation for this cohort. The extent of underestimation should not be severe. Most Canadians who attend university first enroll directly from high school. A survey of students from 26 Canadian Universities suggests that almost 90 per cent of first year undergraduates are younger than 20 years old (Canadian Undergraduate Survey Consortium 2001).[10]

[9] While the YITS project also includes science and math skills tests, I use reading scores because the whole YITS sample wrote the reading test. The sample within each school was randomly divided so that half the students wrote the math and the other half wrote the science test.

[10] Compared to other provinces, some students in Ontario might be slightly older when they first attend university. Until the 2002 academic year students in Ontario might complete five years of high school. In a later section, I show that the results are similar when Ontario is excluded from the

The paper focuses on university participation rather than participation in post-secondary education for a number of reasons; earnings among university graduates are higher than among graduates of other levels of PSE, and because university education is more likely to be affected by factors such as peers and role models. Although estimating the returns to education is difficult, there is evidence that university graduates earn more than graduates of high school or other forms of PSE. Using data from the National Graduates Survey, Drewes (2006) estimated that in 2002, female university graduates earned 34.9 per cent more than college graduates, and male university graduates had 26 per cent higher earnings than college graduates. Ferrer and Riddell (2002) also found evidence of substantial earnings premia in Canadian Census data. Canadian males with university degrees earned 22.8 percent more than males with only a high school diploma. In comparison, college educated males earned 6.6 per cent more than high school graduates. For females, the university premium was 25.2 per cent compared to the college premium of 5.9 per cent.

Another reason to study university participation specifically is that university education is relatively scarce. The majority of Canadians do not have a university degree. In the 2006 Census, less than 25 per cent of Canadians aged 25 to 65 held a Bachelors level degree or higher. The fact that university credentials are not ubiquitous in Canadian society makes it more likely that there are youth who lack information about the costs and benefits. Moreover, attending university often requires leaving one's community. Nearly 97 per cent of Canadian high school students live within commuting distance of a college, whereas only 80 per cent lived near university (Frenette 2003). On the one hand, youth may be more reluctant to attend university if it requires leaving their peer group in their community. On the other hand, youth may follow their peers if those peers are attending university. In contrast, because colleges are available in most communities, the decision to pursue this level of education might be less affected by peer behaviour.

sample.

3.3.2 Control Variables

The YITS provides a data to measure number of important factors that have not been available in other studies or neighbourhood and peer effects, such as reading skills and parents' hopes for their children's future education. All of the control variables used in this paper are measured when the youth are age 15. Table 3.1 reports the means and standard errors for the control variables used in this paper, for the full sample and separately by gender.

Parent's education

Measures of parents' education are derived from data provided by the responding parent in the parent questionnaire. Parents were asked about their own and their spouses' educational attainment. The levels of parental education are collected into four categories according to the highest level of education in the family: less than high school, high school, postsecondary education (PSE) below the Bachelors level, and Bachelors and higher. Less than high school is the reference group. Dummy variables are used for the high school and PSE below BA categories. For families with BAs, rather than using a dummy, I created a variable called family BA share, which is the fraction of parents who have a Bachelors degree. This variable is helpful to gauge the relative importance of the neighbourhood BA share effect.[11]

Adult Equivalent Family Income

Family income combines the respondent's and his or her spouse's before-tax income. The information was collected from the responding parent. The measure is calculated by Statistics Canada and may include imputed values.[12] Income is

[11] The results are not sensitive to this specification. Previous versions have included a set of fully interacted dummies for parents' education. The estimated neighbourhood BA share effect was similar in size and significance.

[12] Total income is derived from a sum of the nine income sources collected during the Parent interview. They are: (1) Wages and Salaries before deductions, including bonuses, tips and commissions; (2) Net Income from Farm and Non-farm Self-employment (after expenses and before taxes); (3) Employment Insurance benefits (before deductions); (4) Canada Child Tax Benefits and provincial child tax benefits or credits (including Quebec Family Allowance); (5) Social Assistance (welfare) and Provincial Income Supplements; (6) Support payments received, such as spousal and child support; (7) Other Government Sources, such as Canada or Quebec Pension Plan Benefits, Old Age Security Pension, or Workers' Compensation Benefits; (8) Goods and Services Tax Credit / Harmonized Tax

then divided by the square root of the number of people living in the household to provide a measure of income per equivalent adult. Because income is unlikely to be linearly related to university participation, I use quartiles of the adult equivalent family income.

Socio-demographic variables

The socio-demographic variables capture family structure, parental labour force status, rural residence, mobility, Aboriginal and immigrant status.

The family structure categories are constructed using the roster of family members provided by the responding parent. The reference category is families with a birth mother and a birth father. Lone mothers are families headed by only the birth mother. Similarly, lone father families are headed by only the birth father. Other two parent families include adoptive parents, combinations of birth and adoptive parents, foster and step parents, and same gendered parents.[13] Other one parent families are those families who are headed by one adult who is not a birth parent of the child.[14]

Parental labour force status is measured by two dummy variables, which take on the value one if either the youth's mother or father works. The variables are zero otherwise.

Socio-demographic background also includes a variable to indicate whether the child's residence when age 15 is in a rural area.[15] Mobility is measured by the number of household moves the child had experienced by age 15.

I include variables indicating that the youth was born in another country and if

Credit received in 1999; and (9) Other Non-Government sources including

[13]Specifically this category includes family compositions of biological mother and stepfather; biological father and stepmother; adoptive mother and adoptive father; biological mother and adoptive father; biological father and adoptive mother; biological mother and foster father; biological father and foster mother; adoptive mother and stepfather; adoptive father and stepmother; stepmother and stepfather; foster mother and foster father; two other guardians (may include same sex partner); biological parent with spouse/partner (may include same sex partner); adoptive / step / foster parent with spouse/partner (may include same sex partner).

[14]Specifically, this category includes family compositions of adoptive/step/foster mother and no father; adoptive/step/foster father and no mother; single guardian (other related or other unrelated); biological parent of unknown gender; adoptive / step / foster parent of unknown gender.

[15]A rural area is defined in Statistics Canada's Statistical Area Classification as population density less than 400 people per square kilometer.

the youth was born in Canada but at least one of her parents were born outside of Canada. I also include a dummy if the youth first learned a non-official language. Finally, I use an indicator for the youths' Aboriginal status. Aboriginal status is reported by the parents and includes North American Indians, Métis and Inuit.

PISA Reading Scores

In the first wave of the survey, respondents completed a PISA reading test. To allow for non-linearity, I use quartiles of the plausible values measure of the reading scores. One should think of this variable as a measure of the *stock* of cognitive skills at age 15. This should include any innate or inherited ability plus the cumulative effect of all the cognitive inputs over the youth's first 15 years. If factors such as family resources, attitudes to education, or neighbourhood quality affect reading skills, then the PISA scores should capture any changes or shocks in these measures that have occurred prior to age 15.

Parents' Hopes for Children's Education

The responding parent is asked to indicate the highest level of education they 'hope' their child will obtain. The responses were coded into four categories, high school or less, PSE below a BA, which includes college diplomas or trade certificates, Bachelors degree or above, and any level above high school. This variable can be used to proxy for a preference for education. The variable may also pick-up information the parent has about the child's ability and motivation.

3.3.3 Neighbourhood Characteristics

A fundamental issue associated with studying neighbourhoods is the difficulty in defining and describing them. This paper uses the fraction of adults who have obtained a Bachelors degree to measure neighbourhood quality.[16] Because of the strong relationship between parental and children's education, this measure of neighbourhood quality might reflect the youths' peer environment. It also describes

[16]In a previous version, a number of other variables including income, population, the unemployment rate and the fraction of immigrants were also used to describe neighbourhoods but these variables were substantially less important. The predicted probability of university participation conditional on the BA share alone is highly correlated (.96) with the predicated probability of attendance conditional on BA share and income, population, the unemployment rate and immigrant share.

the potential non-familial adult role models available to youth.

The youths' neighbourhoods are their home residences measured when they are age 15 and are geographically defined by the Enumeration Area (EA). The EA is the smallest geographic unit used by Statistics Canada and are roughly one to four blocks in size.[17] One benefit of using the EA unit is that the entire territory of Canada is divided into DAs. In contrast, only large metropolitan areas are divided into census tracts.

Neighbourhood characteristics are taken from public use files of the 2001 Canadian Census Profiles. The EA unit was dropped by Statistics Canada between the 1996 and 2001 Census' in favour of the Dissemination Area (DA), which while sometimes smaller than the EA is roughly similar.[18] The EA locations in the YITS were matched to the DA Census 2001 data by adding together the DAs that overlapped within an EA. In other words, if the youth's home EA matched to more than one DA, her neighbourhood was defined as the total area covered by all the matched DAs.

The mean BA share in the data is .158 with a standard deviation of .115. Because DAs are small geographies and the education questions are on the long form of the Census (20 per cent sample), the sample sizes used to estimate the BA share in a neighbourhood can be, in some cases, fairly small. The mean sample size across the YITS-sample neighbourhoods is 196. To gauge the extent to which this might affect the estimates in this paper a sampling variance was calculated as $(1-p)*p/N$, where p is the mean BA share and N is the mean sample size used to estimate the BA share. The sampling deviation is .026 which suggests about 23 per cent of the neighbourhood variation can be attributed to measurement error.

Identification in this paper relies on neighbourhood variation *within* schools. There are 938 schools in the sample; some have very small sample sizes while others have samples over 100 students. The mode sample size is in the order of

[17] Because the boundaries of EAs include any adjacent or surrounded bodies of water, the measures of land area are not available.

[18] The EA was originally designed as the area that a single Census taker could enumerate in a single day. The switch between EAs and DAs was motivated by a change in technology from manually and electronically generated boundaries.

25 students. The median number of neighbourhoods within a school is 9. In the lowest quartile, schools have between 2 and 7 neighbourhoods. Schools with one neighbourhood are not included in the analysis because of lack of variation.

Not surprisingly, variance in neighbourhood BA share is larger between schools than within schools. The between school standard deviation is .09 and the within school standard deviation is .073. In the schools with the least variation in neighbourhood BA share, the standard deviation is .04. At the median the standard deviation is .057 and at the top quartile the standard deviation is .089. Schools explain 62 per cent of the variation in neighbourhood BA share.[19]

3.4 Results

The first column of Table 3.2 shows the results from a Probit regression of university participation on the proportion with a bachelors degree in an Enumeration Area (EA) and gender. The mean of the marginal effects are reported. In other words, I calculated a marginal effect for each sample member and then report the sample average. In all of the models estimated in the paper, the standard errors are clustered by school. Clustering at the school level accounts for any heterogeneity across individuals or neighbourhoods in the school effect. This first specification which demonstrates the degree of correlation between neighbourhoods and individual outcomes can be thought of as a 'base' model to be compared with other specifications.

Because the neighbourhood BA share is measured as a proportion, the coefficient in the first column can be interpreted as percentage points. In other words, the marginal effect reported in column 1 implies that a one percentage point increase in the BA share is associated with .7 of a percentage point increase in the probability of university participation.

Fixed effects for high schools are added in the second column of Table 3.3. While schools may vary in their educational quality, having a direct impact on university participation, they also play an important role in identifying neighbourhood

[19]This estimate comes from the R^2 from a regression of BA share on the school dummies.

effects. A key claim underlying the identification strategy is that school choice is a large part of parents' residential location decisions. If this is true and families sort into schools based on factors such as their children's ability, skills or motivation then adding the school fixed effects should reduce bias from endogenous neighbourhood selection.

After accounting for the youths' high school, the estimated effect of neighbourhood quality drops by roughly 10 percentage points. Moreover, the school fixed effects are jointly significant at the 1 per cent level. This suggests that school fixed effects do control for important factors which are correlated with neighbourhoods and affect university participation.

For comparison, the model reported in the third column includes parental education but no school controls. These effects are relative to a family where both parents were high school drop-outs. It is a well established result that parental education is strongly associated with university participation (Carneiro and Heckman 2002, Black et al. 2005, Oreopoulos et al. 2006, Drolet 2005). Indeed, a youth whose parents both have a BA is 45 percentage points, or nearly three times, more likely to attend university when compared to a youth whose parents are high school drop-outs.

When compared to column 2 with school fixed-effects, the neighbourhood effect falls by a similar magnitude when parental education is added to the model. This is unsurprising since parents are quite likely to sort based on observable factors, such as education. In fact, parental education may be highly correlated with school choice. The school dummies may just proxy for sorting based on parental education. The results in column 4, suggests that schools capture more than just parental education. When school fixed effects and parental education are both added to the model, the neighbourhood effect falls by over half from the base.

Family income quartiles are added in the final column of Table 3.3. Without other socio-economic variables in the model, income has a considerable impact on the probability of attending university. Youth from families with income in the highest quartile are 10 percentage points more likely to attend university relative to the lowest income families. Interestingly, the addition of family income does relatively little to change the neighbourhood effect. This was also case in a specification, not shown, which included only family income and the female dummy.

If families do sort based on their income, it is probably their lifetime income or wealth which matters. Income as measured in surveys, such as the YITS, is generally a poor measure of lifetime income. Table 3.3 adds to the model a set of background characteristics which characterize families' economic resources more fully than spot measures of income. These variables also measure non-monetary background and environmental factors that would affect net utility from university participation. For reference, the first column repeats the results from column 5 in Table 3.2. The second column reports results from a model with measures of socio-demographic background including Aboriginal status, immigrant status, rural residence, household moves, parental labour force status and family structure.

The direct effects of these factors are similar to what one would expect based on previous research on university attendance (Boniskowa 2005, Christofides et al. 2001, Corak et al. 2004, Drolet 2005, Finnie et al. 2004). Native youth are less likely and second generation youth are more likely to attend university. Youth who first spoke a non-official language are also more likely to attend university. This variable captures to a limited extent differences in the source country of immigration.[20]

Instability in childhood is also negatively related to university participation. Each household move a youth experiences is associated with a one percentage point decline in the chance of attending university. Youth from 'other' two parent families are 14 percentage points less likely to be university participants compared to youth who live with both of their biological parents. Other two parent families measure instability because these types of families include custodial guardians and step-parents. After controlling for the demographic characteristics, the estimated

[20]Borjas (1995) has suggested that the ethnic composition of neighbourhoods can influence educational attainment. I do not explore differences in neighbourhood effects by ethnicity in this paper because the data available are not particularly well suited for such analysis. Ethnicity is only available for immigrant and second generation children, who form only a small part of the sample. Later, I demonstrate that the results do not depend on immigrants or second generation children in the sample.

neighbourhood effect falls but by only one tenth of a standard deviation.

The third column of Table 3.3 reports the results from a model where the quartiles of the youths' scores on the PISA reading test are included. The PISA scores measure the stock of reading skills at age 15 and as such are a very important determinant of university participation. Youth who score in the highest quartile are 67 percentage points more likely to attend university compared to those who score in the lowest quartile. Taking reading skills into account also lowers the partial effect of gender, family income and parental education, suggesting that these factors work to some extent through early skills development.

Parental aspirations for their children's educational attainment are included in the specification reported in column 4. These variables should capture parents' preferences for education or any knowledge they have about the abilities and motivations of their children. The probability that a youth attends university is about 39 percentage points higher if her parent hopes she will obtain at least a Bachelors degree, relative to a youth whose parents expect her to achieve high school graduation or less. Within families who hope their children obtain any level of PSE, youth are 30 percentage points more likely to attend university. The youth whose parents expected them to obtain a college or trade certificate were 15 percentage points more likely to attend university than those youth whose parents expected high school or less.

Including parental aspirations in the model once again reduces the neighbourhood effect by 40 per cent of a standard deviation.[21] In this model, which is the specification used throughout the paper, when neighbourhood BA share increases by 10 percentage points, the probability of university attendance increases by 2.5 percentage points, which is roughly one quarter of a standard deviation. This result is considerably smaller than effects found by others. For example, Goux and

[21] Parental aspirations may be a channel through which neighbours impact young people. For example, parents may learn about the value of education from their neighbours and as a result might develop higher expectations for their children's educational attainment. If this was a case, then controlling for parental aspirations would tend to understate the neighbourhood effect. Despite the possibility that neighbourhoods impact youth through parental aspirations I include this control in my main specification because evidence reported in Chapter 2 suggests that measures of parental aspirations are correlated with parental preferences for education.

Maurin (2007), who found that a one standard deviation change in neighbourhood quality resulted in .7 of a standard deviation change in the chances that a youth repeats a grade. The effect is however comparable to the direct effect of parents education. Relative to a youth whose parents are high school drop-outs, a youth with BA educated parents is 32 percentage points more likely to attend university. This compares to the neighbourhood effect of 24 percentage points associated with living in a community where all adults have a BA relative to one where no adults have a BA.

In order to judge whether the school fixed effects continue to capture any factors correlated with neighbourhood quality, the final column of Table 3.3 reports results from a model similar to the column 4 specification but without school fixed effects. Without the school controls, the coefficient is .281 about 4 tenths of a standard deviation higher than the model with school fixed effects.

3.4.1 Sensitivity of Results

This section reports estimates using alternative specifications and different sample restrictions. These robustness checks are intended to gauge how sensitive the results are to factors that might undermine the identifying assumption. The fact that results change very little can help mitigate the concern that the estimated relationship between university participation and neighbourhoods is spurious.

One might be concerned that I find a significant neighbourhood effect simply because there is a correlation between unobserved parents' preferences and neighbourhood characteristics. Although preferences can not be directly observed, the YITS data provides several correlated proxy measures. One particularly helpful measure is the response parents give when questioned about the highest level of education they hope their child obtains. In Table 3.4, I demonstrate that the neighbourhood effects do not depend on parents who want their children to attend university. Parents who have high educational aspirations for their children might be willing to forgo consumption to live near educated families. The first column of Table 3.4 repeats the results from column 4 of Table 3.3 for comparison. The second column of Table 3.4 shows the results from a regression excluding all the children whose parents wanted them to obtain at least a Bachelors degree. Within the sample of youth whose parents have low educational expectations the neighbourhood effect is larger, a ten percentage point change in the neighbourhood BA share increases the chance of attending university by 4.37.

In the third and fourth columns of Table 3.4, I include measures of family behaviour which are correlated with unobserved ability and preferences. When the youth were 15 years old, their parents were asked whether they had done anything specific to ensure that their child would have money for further education after high school. The types of actions ranged from forms of savings, getting additional jobs or encouraging their child to earn money through employment or scholarships.[22] Various ways of saving were the most frequently cited actions taken to ensure children had money for education included opening savings accounts, investing in mutual bonds and starting RESPs.[23] Families who had taken no action, not even as minimal as encouraging their child to save money, in all likelihood, are either unconcerned or believe that their children will not or can not pursue any PSE.

The majority of families did indicate that they had done something to ensure their child had money for PSE-roughly 67 per cent. As one might expect, this variable is correlated with parental aspirations; however, it appears to capture something independent from parents' aspirations. Among parents who hope their child obtains high school or less, 38 per cent had taken some action to make sure their children had money for PSE. About 71 per cent of parents who expected their child to obtain a BA or more had also done something toward financing PSE for their child.

Parents who had not taken any action toward ensuring there was enough money

[22]The question parents were asked is "Have you (or your partner) done anything specific to ensure that child will have money for further education after high school?" Responses to the follow-up asking what the family had done included, started a savings account, started a Registered Education Savings Plan (RESP), set up a trust fund for this child, made investments, such as mutual funds or Canada Savings Bonds, started working or took an additional job, encouraged child to earn money/get a job, encouraged child to work toward a scholarship.

[23]Milligan (2005) finds that variables, such as income and education, that are correlated with educational aspirations are also correlated with investments in RESPs.

for their children's education are also highly unlikely to have specifically chosen their neighbourhood because its characteristics would improve their children's educational outcomes. If neighbourhood BA share only captures the willingness of parents to locate in good neighbourhoods for their children then including this variable should cause the estimated effect to fall substantially.

As Table 3.4 column 3 shows, the youth from families who had taken some action were about 8 percentage points more likely to have attended university. However, when this measure of family's education financing behaviour is included in the regression the estimated neighbourhood effect changes by just over one percentage point.

The behaviour of siblings is another way to measure the unobserved features of a family's propensity for university. Although siblings may experience individual shocks that are orthogonal to their family membership, controlling for sibling behaviour is somewhat like a family effect. Unfortunately, the YITS does not collect information about sibling participation in PSE; however, in cycle one, parents were asked whether any of the youths' siblings had dropped out of high school. This variable along with a dummy variable indicating that the child had no siblings are included in the regression reported in column 4 of Table 3.4.

Youth with a high school dropout sibling are about 12 percentage points less likely to attend university compared to those whose siblings had not dropped out. Including the measure of sibling drop-outs again has little impact on the estimated neighbourhood effect.

One might also be sceptical of the findings if the neighbourhood effect depended on particular sample members. In Table 3.5 the results are reported when various potentially problematic groups are omitted from the sample. The first column of Table 3.5 reports results from a regression where all of the students in Ontario are dropped from the sample. In Ontario, a high school reform that essentially shortened the high school program from 4 to 5 years took effect in 2003. This meant that the cohort born in 1984 graduated from high school at the same time as the cohort born in 1985 and as a result faced increased competition for places in universities. For this reason, I drop students who went to high school in Ontario from the sample. Without Ontario students, the estimated neighbourhood effect is .26, roughly the same as the estimate from the full sample.

In the second column immigrants and second generation youth are dropped from the sample. The relationship between family background and educational attainment is very different for immigrants and second generation youth (Boniskowa 2005). Moreover, sorting into neighbourhoods may have less to do with school choice and more to do with the ethnic enclaves (Borjas 1995). Without immigrants and second generation youth in the sample the estimated impact of neighbourhoods increases by about 8 percentage points. A percentage point change in neighbourhood BA share is associated with one third of percentage point change in the university participation rate.

Because the relationship between school choice and neighbourhood also differs for individuals who attend private schools, this group is omitted from the sample in column 3. Without private schools in the sample, the neighbourhood effect is .301, again larger than the estimate including private schools.

The identification strategy assumes that particular neighbourhoods imply specific schools. At some schools, however, location is not the primary determinant for enrollment. This is the case at most private schools, but also some public schools. Some schools have special arts or sports programs, for example. The model in column 4 includes only schools where the principal reported enrollment was always determined by location of residence. In this sample, the estimated effect is .295.

Enrolling their children in French immersion is another reason why families might not view their choice of home location as tied to their children's school. French immersion is very desirable in Canada in part because bilingualism is valued and also in part because parents believe immersion programs deliver higher quality with respect to curriculum and the children's peers. Not every high school has an immersion program so effectively immersion programs have large catchment areas which weakens the argument supporting my identification. For this reason, I drop all the children in French immersion in the model reported in the final column of Table 3.5.

The last column of Table 3.5 shows the results using a sample where families who indicated that they hoped their child would attain a university degree are excluded. These youth are excluded because one might expect that among them are families with extreme preferences for education who might be willing to wait for houses to come available in neighbourhoods with a high BA share. In the sample without families with BA aspirations for their children, the neighbourhood effect is quite large, about .44 of a percentage point.

Finally, in Table 3.6, I check whether the relationship between university participation and neighbourhoods is arbitrary. Although I have not tried to claim that the estimates identify a specific effect, such as peer effects, I do interpret these effects as being related to social interaction. It is important therefore to demonstrate that the relationship between BA share and university participation captures some correlation which is not entirely spurious.

In Table 3.6, I estimate the effect of alternative measures of neighbourhood on university participation. In column 1, instead of BA share, I use the fraction of dwellings in the neighbourhood that were built between 1996 and 2001. In the second column, I include both the age of housing stock measure and the BA share. The fraction of dwellings that are single detached homes is used in the third column, and in combination with the BA share in the last column. New houses and detached homes are desirable among families and as such might be correlated with housing demand. Housing types should not, however, directly affect university participation. Evidence of a significant relationship between these measures of neighbourhood quality and university participation would suggest that the relationship between neighbourhood BA share and participation is spurious. As Table 3.6 shows there is no such evidence. Both measures of neighbourhood quality are not related to university participation. The estimates are small and statistically insignificant. Their inclusion also does not qualitatively change the estimated neighbourhood effects.

3.4.2 Marginal Effects Across the Socio-Economic Distribution

The probability of university participation varies significantly among youth from different socio-economic (SES) backgrounds, which reflects underlying differences in the net utility from university. For some youth, the decision to attend university is obvious. For example, some youth may struggle to pass high school courses and would not consider university an option. Others may excel in high school, and may have considerable encouragement and financial support from their parents, which might make attending university automatic. For these youth, neighbourhood influences are unlikely to change the outcome of university participation.

I test for significant differences in the coefficients across subgroups defined by parental education, PISA scores and parents' hopes for their children's educational attainment.[24] The estimated coefficients are reported in Table 3.7. These results suggest that the coefficients on neighbourhood BA share do not differ statistically by PISA scores or parents' aspirations for their children's education. There are, however, statistically significant differences in the coefficients across parents' education categories. It is much easier to understand the implications of differences in estimated coefficients by examining the resulting differences in marginal effects.

Differences across the SES distribution in the marginal effects of neighbourhoods on university participation stem in part from differences in the coefficients and the density associated with the particular SES characteristics. The rest of this section reports estimates of the marginal effect of neighbourhoods on the probability of university attendance at different points on the SES distribution. SES types are defined according to the youths' family income, their parents' education, their PISA scores and what level of education their parents hoped they would achieve. Four points on the SES distribution, called "Low", "Medium-High School", "Medium-PSE", and "High", are defined in Table 3.8. All the other variables in the model are evaluated at their mean, throughout this section.

The top panel of Figure 3.1 reports the estimated marginal effects and 95 % confidence intervals by SES type separately for boys and girls.[25] The estimates shown correspond to the percentage point change in the probability of university attendance resulting from a 10 percentage point change in neighbourhood BA share,

[24]Interactions between gender and family income are not reported here because they were statistically and economically insignificant in a wide variety of specifications.

[25]The results are shown separately by gender because of the considerable difference in participation rates. The average girl has a higher PISA score relative to the average boy.

which is roughly one standard deviation. The bottom panel reports the predicted probability of attendance in percent.

The strong relationship between SES, skills and university participation is immediately evident from the predicted probabilities in Figure 3.1. The probability of attendance for a High type girl is 95 percentage points higher than a Low type. High types are youth who scored in the highest quartile of the PISA distribution, and whose parents have a BA, and hoped their child would achieve a BA, and who have family income in the top quartile. In contrast, Low types are drawn from the lowest skills quartile, are in the lowest family income quartile, and have parents who are high school dropouts and hop their child will achieve a high school diploma. For boys, the difference in university participation rates between High and Low types is just as striking. While less than one per cent of Low type boys are predicted to attend university, High type boys are well over 100 times more likely to be university participants. In other words, low- skilled youth from disadvantaged families have virtually no chance of attending university, while high-skilled youth from advantaged families are almost guaranteed to attend university. Another notable feature of the distribution of predicted probabilities is the steepness of the gradient. While the predicted probability for each SES type is relatively larger, the probability of attendance is above the mean only for High types.

The marginal effects shown in the top panel of Figure 3.1 create an inverted U shape across the distribution of SES types. While the shape itself is simply a feature of the Probit estimation, what is interesting is for which type the marginal effect is largest.[26] For both boys and girls, neighbourhoods have the largest impact for the Medium-PSE type. A 10 percentage point increase in BA share is associated with a 1.92 percentage point increase in the chances that Medium-PSE type boys participate in university and a 2.25 percentage point increase in the participation rate among girls. The marginal effects are statistically similar across the three lower SES categories. The High type is the only type for which the marginal effect is statistically significantly different from that of Medium-PSE types, for both boys

[26]The partial effect estimated in a Probit is proportional to the normal density evaluated at the particular characteristics vector of interest.

and girls.

One way to interpret the results in Figure 3.1 is that the marginal effect of neighbourhoods are largest for types who are closest to the margin of university participation. For example, Medium-High School types are closer to the margin of university participation relative to High types, for whom the marginal effect of neighbourhoods is zero. With that interpretation in mind, I next explore how the marginal effects change as I shift one component of SES type at a time.

Figures 3.2 and 3.3 show the marginal effects and predicted participation rates among boys and girls for each SES type evaluated at different PISA quartiles. Even within a PISA reading skills quartile, family background still has a large effect on the chances that a youth attends university. For boys, the participation rate for Low types who are drawn from the top of the skills distribution is roughly 25 per cent, below the average of 33 per cent. The Low type is a youth whose parents are dropouts have the lowest quartile of income and hope their child graduates from high school. This compares to 93 per cent for similarly skilled High types. These are youth whose parents have a BA, fall within the highest quartile of income and hope their child achieves a university degree.

If the PISA reading test measures something related to ability or skills that is valuable in the market for university educated workers then policy makers should be interested in ensuring that those drawn from the top of the PISA distribution are well represented among university participants. The results in Figures 3.2 and 3.3 suggest that for highly skilled youth from less advantaged backgrounds, university participation is a decision that is influenced by their neighbourhoods. The same cannot be said of highly skilled youth from advantaged families.

For the two lowest SES types, marginal neighbourhood effects reach a maximum at the top quartile of PISA scores. Independent of skills, the marginal neighbourhood effect on youth from families with BAs is statistically insignificant and economically small. Indeed, the estimated effect for both boys and girls is negative and less than one half of a percentage point for a 10 percentage point change in BA share. This compares to a marginal effect for high-skilled youth whose parents are high school graduates, have income in the second quartile and hope their child achieves some level of PSE of nearly 8 percentage points for girls and 6.5 percentage points for boys, for a 10 percentage point change in neighbourhood BA share.

The estimates for low-skilled youth from disadvantaged families are very small, which is consistent with the literature. In contrast, it may seem surprising that the effects for highly skilled youth also from disadvantaged families are so large. Put in context, however, these marginal effects may not seem implausibly large. In the data, over 70 per cent of all youth in the highest skills quartile attend university. For highly skilled youth from Low type families, the predicted probability of university attendance is .25 for boys, and .38 for girls. A two standard deviation increase in the neighbourhood BA share is not large enough to bring highly skilled youth from disadvantaged families up to the average probability of participation among similarly skilled youth.

Another reason to have confidence in the results in Figures 3.2 and 3.3 is that the distribution of PISA scores is not correlated with variation in residual BA share conditional on SES background and school selection.[27] If most of the highly-skilled youth from disadvantaged backgrounds were clustered in more educated neighborhoods, one might suspect that some family characteristic, such as parents' preferences for education was biasing the results. However, this does not appear to be the case. After conditioning on schools, family background, and the parents' reported aspirations for their child's education, neighbourhoods bear no statistical relationship to PISA scores. The mean PISA score in the highest quartile of residual neighbourhood BA share is 542.83 compared to the mean in the lowest quartile of 541.31, a difference which is not statistically significant.

In the next set of figures, I now consider how the marginal effects differ as I shift the family income quartiles for the SES types. Because only the coefficient for the top quartile of family income was statically significant, it is not particularly surprising to see in Figures 3.4 and 3.5 that the marginal effect of neighbourhoods

[27] Residual BA share is estimated from regressing neighbourhood BA share on parents' education, family income, parents' educational aspirations, socio-demographic characteristics and school fixed effects. Socio-demographic characteristics include family structure, parents' labour force status, youth's Aboriginal and immigrant status, the number of household moves, and rural residence.

does not vary across income quartiles for either boys or girls. It is important to interpret these findings in the context of the Canadian policy environment. Tuition at Canadian universities is arguably low relative to the United States and a fairly generous student loans program is available in Canada for low-income students.[28] In the absence of these policies, it is possible that income might play a much larger role.

Finally, in Figures 3.6 and 3.7, I shift from parents who hope their children attain some PSE to parents who hope their children attain a university degree. For the two lowest SES types, marginal neighbourhood effects are larger when parents hope their children will attain a university education. For youth drawn from the third quartile of the skills distribution and whose parents have non-university PSE credentials and are in the third quartile of income, the difference across parental aspirations in the marginal effects is not statistically significant. The same is true of the most advantaged youth, that is the High types. For this group, the neighbourhood effect is always zero.

3.4.3 The Results in Comparison to Other Studies

The findings in this paper are consistent with the literature which finds that neighborhood quality has had little or no impact on economic outcomes for individuals drawn from the bottom of the socio-economic distribution. Although the marginal effect of neighbourhood quality was statistically significant for low-SES youth, the size of the effect was very small and compares to the size of effects found using data from similar populations. For example, the Moving to Opportunity experiment did not statistically significantly affect university participation among the children of participants. In the MTO control group, about 4.3 per cent of participants' children attended university five years after the experiment began. This compared to 4.8 per cent in the experiment group. The difference in neighbourhood quality for these

[28]It is difficult to make tuition comparisons because quality may vary substantially across different types of institutions in Canada and the United States. The average full-time undergraduate tuition across Canadian universities was $4,214 in 2007-08 (Canada 2005). In 2005-2006, the average tuition for undergraduates at U.S. universities was $11,338. At U.S. public universities the average in-state tuition is $6,836. (National Center for Education Statistics 2007).

participants is about a 2 percentage point improvement in the share of families in the neighbourhood with a college degree (Orr et al. 2003). My results suggest that for families with less than a high school diploma a 2 percentage point increase in the neighbourhood BA share would be associated with a difference in university participation of 9 one-hundredths of one percentage point. In other words, the neighbourhood effects estimated for a similar population is actually much smaller than the effect estimated in the MTO experiment.

Among the studies that have found significant impacts of neighbourhoods, the size of various effects range from modest to very large. Direct comparisons are difficult because studies use different neighbourhood characteristics and consider different outcomes. The studies with large effects are typically focused on health and crime outcomes. Datcher (1982) is one paper to which a relevant comparison can be made. This paper, which uses the Michigan Panel Study of Income Dynamics, is particularly interesting because it is a rare example that uses a measure of parental aspirations for children's education as a control variable. Datcher (1982) regresses years of education on mean neighbourhood income so the exact size of the effect is not comparable. However, in comparison to similar studies at that time period, the effects are modest. Datcher (1982) finds for white males an increase of $10,000 in mean neighbourhood income is significantly associated with a one year increase in schooling.

3.5 Conclusion

This paper has demonstrated that neighbourhoods are strongly related to university participation. As one would expect, across the socio-economic distribution the marginal effect of neighbourhoods is always substantially smaller than the effect of skills and other important family background characteristics including parents' education, and parental aspirations for their children's educational attainment. At the mean, the estimated effect of living in a neighbourhood where all your neighbours have a Bachelors degree relative to a neighbourhood where none have a BA is about 40 per cent of the impact of scoring in the top quartile on the reading skills test relative to the bottom quartile, and three quarters of the impact of growing up with parents who hold a BA relative to parents who are high school drop-outs.

The marginal effect of neighbourhoods on the probability of university attendance differs substantially by socio-economic background. For youth whose parents have a Bachelors degree, neighbourhoods have no effect on university participation. Youth in the middle of the SES distribution are more likely to be affected by social interactions in their neighbourhood. Neighbourhoods have a statistically significant effect on the marginal probability of attending university among low-skilled youth from disadvantaged backgrounds, but the effect is very small. This finding is consistent with the literature that also finds small or no neighbourhood impacts within disadvantaged populations.

A key contribution of this paper is the analysis of marginal effects across different measures of skills and family background. The paper reveals an interesting interaction between family background, and the pattern of marginal neighbourhood effects along the distribution of reading skills. One of the most striking findings is that while university participation among youth from families with BAs is unaffected by neighbourhoods independent of skills, the marginal effect of neighbourhoods is largest for the most skilled youth whose parents have high school or less, below median income and who have low educational aspirations for their children. This suggests that highly skilled youth are most vulnerable to negative influences in their neighbourhood when they come from disadvantaged backgrounds.

Overall, the results suggest that parents who have Bachelors degrees provide a family environment that orients youth toward university so that independent of their reading skills and influences in their neighbourhood, these youth attend university with above average probability. This conclusion is consistent with Carneiro and Heckman (2002) and the notion of a missing market for family. In other words, the SES gradient in educational attainment exists in part because children can not 'purchase' the family background that leads to university participation. This paper provides some new insights with respect to the connection between family background and university participation. The conclusion in Carneiro and Heckman (2002) implies a remedy of public investments in early childhood development.

Insofar as early childhood development leads to reading skills in adolescence, my results suggest that early childhood investments that bring disadvantaged youth to the same skill level as advantaged youth might not be enough to level the playing field.

The result that neighbourhoods have little effect on disadvantaged youth suggests that relocation policies, or desegregation policies may have limited impact on some labour market outcomes where disadvantaged individuals are far from the margin. There may be other outcomes, however, where disadvantaged individuals are sufficiently near the margin so that peer and role model effects do have a meaningful impact. These could include choices about health or parenting styles, and early investments in children's development.

Because of the emphasis on marginal youth, one lesson that can be taken away from the results is the importance of designing policies that can be targeted at specific populations who are most at risk. Rather than broad based relocation policies, it might be the case that smaller scale mentoring type programs are more appropriate. In this context, it might be quite difficult given the types of policy instruments available to governments to define the relevant subgroups who are most likely to benefit from such programs. Effective programs might define target populations by combining objective characteristics such as family background and referral systems involving teachers or community center leaders who have closer contact with youth and are better placed to identify those on the margins.

Table 3.1: Sample Summary Statistics (Standard Errors in Parenthesis)

	Full Sample	Boys	Girls
University participation	0.411	0.334	0.488
	(0.004)	(0.006)	(0.006)
Female	0.501		
	(0.004)		
Parental Education			
Less than high school	0.234	0.229	0.238
	(0.004)	(0.005)	(0.005)
High school	0.234	0.229	0.238
	(0.004)	(0.005)	(0.005)
PSE below BA	0.454	0.459	0.450
	(0.004)	(0.006)	(0.006)
Family BA share	0.191	0.191	0.191
	(0.003)	(0.004)	(0.004)
Family Income Quartiles			
Q1	0.196	0.175	0.216
	(0.003)	(0.005)	(0.005)
Q2	0.222	0.226	0.219
	(0.004)	(0.005)	(0.005)
Q3	0.277	0.285	0.270
	(0.004)	(0.006)	(0.005)
Q4	0.304	0.314	0.295
	(0.004)	(0.006)	(0.005)
Reading Quartiles			
Q1	0.198	0.258	0.139
	(0.003)	(0.005)	(0.004)
Q2	0.238	0.246	0.230
	(0.004)	(0.005)	(0.005)
Q3	0.271	0.265	0.278
	(0.004)	(0.005)	(0.005)
Q4	0.292	0.231	0.354
	(0.004)	(0.005)	(0.006)
Parents Educational Aspirations			
High school or less	0.026	0.030	0.021
	(0.001)	(0.002)	(0.002)
College or Trade	0.268	0.307	0.229
	(0.004)	(0.006)	(0.005)
BA or above	0.635	0.593	0.678
	(0.004)	(0.006)	(0.006)
Any PSE	0.071	0.071	0.072
	(0.002)	(0.003)	(0.003)
Sample	13,611	6,613	6,998

continued, next page

Table 3.1 cont'd: Sample Summary Statistics (Standard Deviation in Parenthesis)

	Full Sample	Boys	Girls
Aboriginal	0.025 (0.001)	0.027 (0.002)	0.023 (0.002)
Immigrant	0.079 (0.002)	0.073 (0.003)	0.086 (0.003)
Second Generation	0.186 (0.003)	0.185 (0.005)	0.186 (0.005)
Non-official language first spoken	0.111 (0.003)	0.104 (0.004)	0.117 (0.004)
Rural	0.251 (0.004)	0.247 (0.005)	0.254 (0.005)
Number of moves	2.060 (0.021)	2.029 (0.032)	2.091 (0.029)
Mother works	0.761 (0.004)	0.759 (0.005)	0.764 (0.005)
Father Works	0.808 (0.003)	0.827 (0.005)	0.789 (0.005)
Two Parent Family	0.135 (0.003)	0.120 (0.004)	0.150 (0.004)
Lone mother	0.135 (0.003)	0.120 (0.004)	0.150 (0.004)
Lone Father	0.027 (0.001)	0.029 (0.002)	0.025 (0.002)
Other one parent family	0.112 (0.003)	0.107 (0.004)	0.117 (0.004)
Other two parent family	0.112 (0.003)	0.107 (0.004)	0.117 (0.004)
Sample	13,611	6,613	6,998

Source: Youth in Transition Survey–Cohort A

Table 3.2: Impact of Neighbourhoods on University Participation

	1	2	3	4	5
Neighbourhood BA share	0.704 (0.033)***	0.595 (0.069)***	0.524 (0.027)***	0.328 (0.072)***	0.278 (0.072)***
Female	0.137 (0.006)***	0.184 (0.010)***	0.110 (0.004)***	0.195 (0.010)***	0.198 (0.010)***
Parental Education–reference both parents < H.S.					
High school			0.037 (0.008)***	0.063 (0.019)***	0.059 (0.019)***
PSE below BA			0.085 (0.006)***	0.168 (0.018)***	0.154 (0.017)***
Family BA share			0.450 (0.011)***	0.545 (0.025)***	0.491 (0.030)***
Family Income Quartiles–reference lowest Q					
Q2					0.019 (0.015)
Q3					0.067 (0.015)***
Q4					0.102 (0.016)***
Predicted probability	.407	.395	.405	.394	.394
Log Likelihood	-5078.8	-4581.6	-4770.8	-4312.4	-4295.2
School F.E.	N	Y	N	Y	Y
Sample	13,611	13,611	13,611	13,611	13,611

Estimates from a Probit model predicting university participation.
Mean of marginal effects are shown (standard errors in parenthesis)
Source: Youth in Transition Survey–Cohort A and 2001 Canadian Census Profiles.
Standard errors clustered by school
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10 level.

Table 3.3: Controlling for Socio-Demographics, Skills and Parent's Educational Aspirations

	1	2	3	4	5
Neighbourhood BA share	0.278 (0.072)***	0.262 (.072)***	0.288 (0.079)***	0.245 (0.077)***	0.281 (0.065)***
Female	0.198 (0.010)***	0.206 (0.010)***	0.137 (0.020)***	0.132 (0.028)***	0.116 (0.011)***
Parental Education–reference both parents $<$ H.S.					
High school	0.059 (0.019)***	0.072 (0.019)***	0.052 (0.023)**	0.051 (0.025)**	0.054 (0.021)***
PSE below BA	0.154 (0.017)***	0.166 (0.019)***	0.105 (0.025)***	0.088 (0.028)***	0.102 (0.020)***
Family BA share	0.491 (0.030)***	0.512 (0.026)***	0.367 (0.028)***	0.321 (0.028)***	0.318 (0.026)***
Family Income Quartiles–reference lowest Q					
Q2	0.019 (0.015)	0.016 (0.015)	-0.009 (0.016)	-0.015 (0.017)	-0.016 (0.016)
Q3	0.067 (0.015)***	0.074 (0.016)***	0.029 (0.018)	0.012 (0.018)	0.005 (0.016)
Q4	0.102 (0.016)***	0.114 (0.018)***	0.074 (0.022)***	0.047 (0.022)**	0.037 (0.017)**
Reading Quartiles–reference lowest Q					
Q2			0.259 (0.038)***	0.221 (0.046)***	0.165 (0.020)***
Q3			0.484 (0.082)***	0.424 (0.086)***	0.331 (0.024)***
Q4			0.669 (0.174)***	0.613 (0.180)***	0.511 (0.034)***
Parents' Educational Aspirations–reference HS or less					
College or trade				0.149 (0.065)**	0.133 (0.066)**
BA or above				0.388 (0.061)***	0.377 (0.054)***
Any PSE				0.303 (0.101)***	0.293 (0.077)***
Predicted Probability	.394	.391	.367	.352	.372
Log Likelihood	-4295.21	-4208.582	-3578.849	-3430.891	-3902.135
School F.E.	Y	Y	Y	Y	N
Sample	13,611	13,611	13,611	13,611	13,611

continued, next page

Table 3.3 cont'd: Controlling for Socio-Demographics, Skills and Parent's Educational Aspirations

	1	2	3	4	5
Aboriginal		-0.070	-0.043	-0.063	-0.027
		(0.029)**	(0.035)	(0.042)	(0.032)
Immigrant		0.100	0.177	0.145	0.088
		(0.033)***	(0.045)***	(0.048)***	(0.034)***
Second generation		0.065	0.089	0.073	0.054
		(0.017)***	(0.023)***	(0.025)***	(0.018)***
Non-official language first spoken		0.143	0.167	0.142	0.104
		(0.028)***	(0.039)***	(0.043)***	(0.027)***
Rural		-0.066	-0.026	-0.019	0.015
		(0.025)***	(0.026)	(0.027)	(0.019)
Number of moves		-0.012	-0.015	-0.016	-0.013
		(0.003)***	(0.003)***	(0.003)***	(0.003)***
Mother works		-0.015	-0.008	-0.008	-0.011
		(0.011)	(0.012)	(0.013)	(0.012)
Father works		0.016	0.006	-0.002	-0.022
		(0.020)	(0.022)	(0.022)	(0.021)
Family Structure–reference two parent family					
Lone mother		-0.027	-0.040	-0.061	-0.086
		(0.024)	(0.027)	(0.034)*	(0.024)***
Lone father		-0.137	-0.137	-0.131	-0.149
		(0.042)***	(0.061)**	(0.078)*	(0.044)***
Other two parent family		-0.130	-0.115	-0.115	-0.107
		(0.019)***	(0.038)***	(0.055)**	(0.021)***
Other one parent family		-0.103	-0.059	-0.095	-0.050
		(0.071)	(0.075)	(0.085)	(0.073)
Predicted Probability	.394	.391	.367	.352	.372
Log Likelihood	-4295.21	-4208.582	-3578.849	-3430.891	-3902.135
School F.E.	Y	Y	Y	Y	N
Sample	13,611	13,611	13,611	13,611	13,611

Estimates from a Probit model predicting university participation.
Mean of marginal effects are shown (standard errors in parenthesis)
Source: Youth in Transition Survey–Cohort A and 2001 Canadian Census Profiles.
Standard errors clustered by school
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10 level.

Table 3.4: Controlling for Proxy Measures of Family Attitudes to Education

	1	2	3	4
Neighbourhood BA share	0.245 (0.077)***	0.437 (0.188)**	0.231 (0.100)**	0.228 (0.100)**
Family prepared to finance education			0.077 (0.016)***	
Sibling dropped out				-0.119 (0.029)***
Only child				-0.026 (0.033)
Female	0.132 (0.028)***	0.126 (0.025)***	0.133 (0.016)***	0.134 (0.016)***
Parental Education–reference both parents < H.S.				
High school	0.051 (0.025)**	0.136 (0.051)***	0.051 (0.034)	0.046 (0.034)
PSE below BA	0.088 (0.028)***	0.110 (0.040)***	0.084 (0.031)***	0.082 (0.031)***
Family BA share	0.321 (0.028)***	0.401 (0.061)***	0.313 (0.043)***	0.313 (0.043)***
Family Income Quartiles–reference lowest Q				
Q2	-0.015 (0.017)	-0.054 (0.035)	-0.021 (0.027)	-0.018 (0.027)
Q3	0.012 (0.018)	-0.018 (0.036)	-0.001 (0.029)	0.011 (0.029)
Q4	0.047 (0.022)**	-0.006 (0.041)	0.031 (0.029)	0.046 (0.029)
Reading Quartiles–reference lowest Q				
Q2	0.221 (0.046)***	0.150 (0.047)***	0.222 (0.030)***	0.224 (0.030)***
Q3	0.424 (0.086)***	0.387 (0.052)***	0.428 (0.026)***	0.426 (0.025)***
Q4	0.613 (0.180)***	0.627 (0.052)***	0.616 (0.022)***	0.614 (0.022)***
Parents Educational Aspirations–reference HS or less				
College or trade	0.149 (0.065)**	0.092 (0.045)*	0.135 (0.060)**	0.144 (0.062)**
BA or above	0.388 (0.061)***		0.375 (0.043)***	0.384 (0.043)***
Any PSE	0.303 (0.101)***	0.305 (0.084)***	0.290 (0.064)***	0.300 (0.065)***
Predicted Probability	0.352	0.154	0.352	0.349
Log Likelihood	-3430.891	-1239.3563	-5855.415	-5839.649
School F.E.	Y	Y	Y	Y
Sample	13,611	3,365	13,595	13,581

Estimates from a Probit model predicting university participation.
Mean of marginal effects are shown (standard errors in parenthesis)
Source: Youth in Transition Survey–Cohort A and 2001 Canadian Census Profiles.
Standard errors clustered by school
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10 level.
All specifications include controls for aboriginal status, immigrant and second generation, non-official language first spoken, rural residence, number of household moves, parents' labour force status, and family structure.

Table 3.5: Alternative Samples

	Without Ontario	Without Immigrants	No Private Schools	Only Res. Schools	Without French Immersion
Neighbourhood BA share	0.259	0.329	0.301	0.295	0.205
	(0.114)**	(0.120)***	(0.105)***	(0.115)**	(0.105)**
Female	0.114	0.136	0.136	0.138	0.133
	(0.016)***	(0.017)***	(0.016)***	(0.018)***	(0.017)***
Parental Education–reference both parents < H.S.					
High school	0.066	0.009	0.063	0.017	0.062
	(0.032)**	(0.036)	(0.035)*	(0.039)	(0.036)*
PSE below BA	0.097	0.058	0.104	0.063	0.089
	(0.028)***	(0.032)*	(0.032)***	(0.036)*	(0.033)***
Family BA share	0.333	0.299	0.347	0.288	0.329
	(0.043)***	(0.046)***	(0.045)***	(0.050)***	(0.046)***
Family Income Quartiles–reference lowest Q					
Q2	0.016	0.003	-0.023	-0.035	-0.016
	(0.026)	(0.028)	(0.027)	(0.030)	(0.029)
Q3	0.038	0.027	0.002	-0.013	0.027
	(0.025)	(0.030)	(0.028)	(0.032)	(0.030)
Q4	0.076	0.073	0.034	0.012	0.053
	(0.026)***	(0.030)**	(0.029)	(0.031)	(0.030)*
Reading Quartiles–reference lowest Q					
Q2	0.188	0.251	0.213	0.203	0.217
	(0.027)***	(0.032)***	(7.030)***	(0.035)***	(0.031)***
Q3	0.374	0.459	0.419	0.422	0.419
	(0.027)***	(0.029)***	(14.940)***	(0.030)***	(0.026)***
Q4	0.565	0.646	0.023	0.606	0.615
	(0.024)***	(0.023)***	(21.190)***	(0.026)***	(0.023)***
Parents Educational Aspirations–reference HS or less					
College or trade	0.132	0.124	0.063	0.094	0.197
	(0.054)**	(0.067)*	(2.130)**	(0.070)	(0.073)***
BA or above	0.362	0.336	0.044	0.353	0.417
	(0.039)***	(0.051)***	(7.170)***	(0.051)***	(0.049)***
Any PSE	0.247	0.274	0.067	0.228	0.326
	(0.060)***	(0.077)***	(4.240)***	(0.078)***	(0.072)***
Predicted Probability	0.331	0.290	0.337	0.338	0.351
Log Likelihood	-5052.3048	-4616.5438	-5545.0377	-4174.5752	-4949.8222
School F.E.	Y	Y	Y	Y	Y
Sample	11,533	11,016	9,760	12,958	11,474

Estimates from a Probit model predicting university participation.
Mean of marginal effects are shown (standard errors in parenthesis)
Source: Youth in Transition Survey–Cohort A and 2001 Canadian Census Profiles.
Standard errors clustered by school
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10 level.
All specifications include controls for aboriginal status, immigrant and second generation, non-official language first spoken, rural residence, number of household moves, parents' labour force status, and family structure.

Table 3.6: Alternative Measures of Neighbourhood

	1	2	3	4
Neighbourhood BA share		0.260 (0.098)***		0.228 (0.099)**
Proportion of dwellings built after 1996	-0.063 (0.069)	-0.086 (0.070)		
Proportion of dwellings that are single detached			-0.150 (0.106)	-0.132 (0.108)
Female	0.129 (0.015)***	0.129 (0.015)***	0.128 (0.015)***	0.128 (0.015)***
Parental Education–reference both parents $<$ H.S.				
Both HS	0.170 (0.057)***	0.172 (0.057)***	0.171 (0.057)***	0.173 (0.057)***
Both PSE below BA	0.246 (0.052)***	0.246 (0.052)***	0.245 (0.053)***	0.246 (0.053)***
Both BA or more	0.465 (0.048)***	0.458 (0.049)***	0.464 (0.049)***	0.459 (0.049)***
Family Income Quartiles–reference lowest Q				
Q2	-0.016 (0.026)	-0.017 (0.026)	-0.015 (0.026)	-0.016 (0.026)
Q3	0.011 (0.028)	0.009 (0.028)	0.012 (0.028)	0.010 (0.028)
Q4	0.040 (0.027)	0.034 (0.027)	0.041 (0.027)	0.035 (0.027)
PISA Quartiles–reference lowest Q				
Q2	0.223 (0.029)***	0.222 (0.029)***	0.222 (0.029)***	0.221 (0.029)***
Q3	0.427 (0.025)***	0.428 (0.025)***	0.427 (0.025)***	0.427 (0.025)***
Q4	0.612 (0.022)***	0.613 (0.022)***	0.612 (0.022)***	0.613 (0.022)***
Parents' Educational Aspirations–reference HS or less				
College or trade	0.139 (0.059)**	0.138 (0.060)**	0.138 (0.059)**	0.137 (0.060)**
BA or above	0.377 (7.660)***	0.375 (7.450)***	0.376 (7.620)***	0.375 (7.440)***
Any PSE	0.283 (0.062)***	0.283 (0.064)***	0.283 (0.062)***	0.283 (0.063)***
Predicted Probability	.352	.352	.352	.352
Log Likelihood	-5887.679	-5880.607	-5885.036	-5879.527
School F.E.	Y	Y	Y	Y
Sample	13,611	13,611	13,611	13,611

Estimates from a Probit model predicting university participation.
Mean of marginal effects are shown (standard errors in parenthesis)
Source: Youth in Transition Survey–Cohort A and 2001 Canadian Census Profiles.
Standard errors clustered by school
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10 level.
All specifications include controls for aboriginal status, immigrant and second generation, non-official language first spoken, rural residence, number of household moves, parents' labour force status, and family structure.

Table 3.7: Heterogenous Neighbourhood Effects

Neighbourhood BA share	2.043 (0.679)***
Parental Education–reference both parents < H.S.	
High school	0.271 (0.091)***
PSE below BA	0.486 (0.084)***
Family BA share	1.274 (0.121)***
Neighbourhood BA Share * Parental Education–reference both parents < H.S.	
Neighbourhood BA Share * High school	-0.740 (0.600)
Neighbourhood BA Share *PSE below BA	-1.570 (0.536)***
Neighbourhood BA Share *Family BA share	-2.310 (0.640)***
Reading Quartiles–reference lowest Q	
Q2	0.579 (0.079)***
Q3	1.078 (0.079)***
Q4	1.738 (0.087)***
Neighbourhood BA Share *Reading Quartiles–reference lowest Q	
Neighbourhood BA Share *Q2	-0.020 (0.492)
Neighbourhood BA Share * Q3	0.344 (0.496)
Neighbourhood BA Share *Q4	-.010 (0.533)
Parents' educational aspirations–reference HS or less	
College or Trade	0.377 (0.143)***
BA or above	1.157 (0.147)***
Any PSE	0.757 (0.147)***
Neighbourhood BA Share *Parents' educational aspirations–reference any level below BA	
Neighbourhood BA Share *BA or above	-0.038 (0.321)
Log Likelihood	-3424.262
School F.E.	Y
Family background controls	Y
Sample	13,611

Estimates from a Probit model predicting university participation.
Coefficients are shown (standard errors clustered by school in parenthesis)
Source: Youth in Transition Survey–Cohort A and 2001 Canadian Census Profiles.
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10 level.
Specification include controls for aboriginal status, immigrant and second generation, non-official language first spoken, rural residence, number of household moves, parents' labour force status, and family structure.

Table 3.8: SES Types Used to Estimate Marginal Effect of Neighbourhood on University Participation

	Low	Medium-HS	Medium-PSE	High
Parental Education	< High school	High school	PSE below BA	BA or higher
Family Income Quartile	1	2	3	4
PISA Quartile	1	2	3	4
Parents' hope for child's education	High school or less	College/Trade or Any PSE	College/Trade or Any PSE	BA or higher

Parents' hope for child's education measures response to the question "What is the highest level of education you hope your child achieves?"

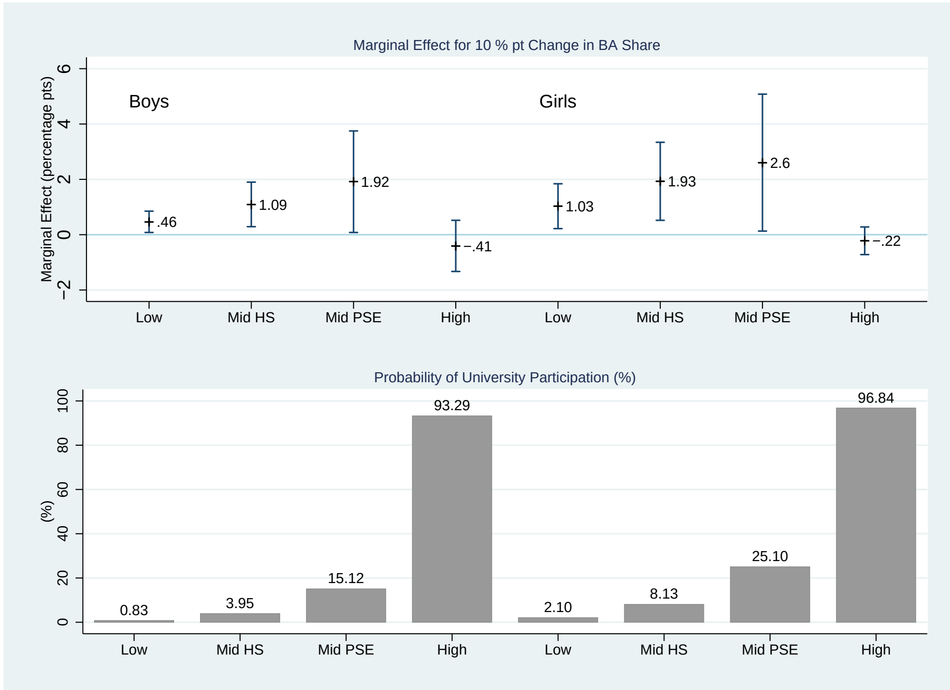

Figure 3.1: Predicted Probability of University Participation and Marginal Neighbourhood Effects by SES Type

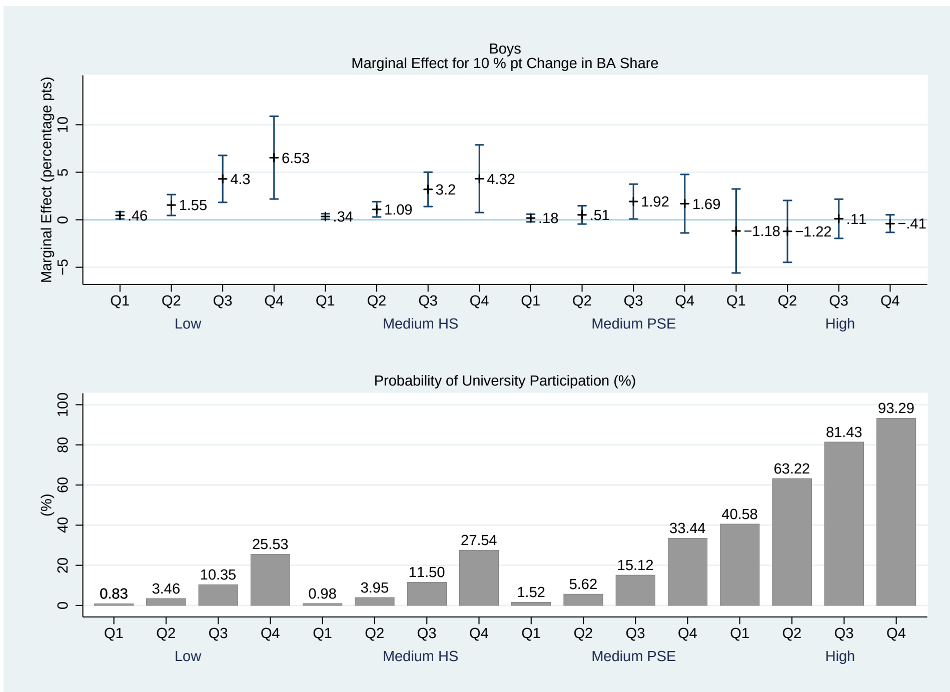

Figure 3.2: Predicted Probability of University Participation and Marginal Neighbourhood Effects by PISA Quartile and SES Type for Boys

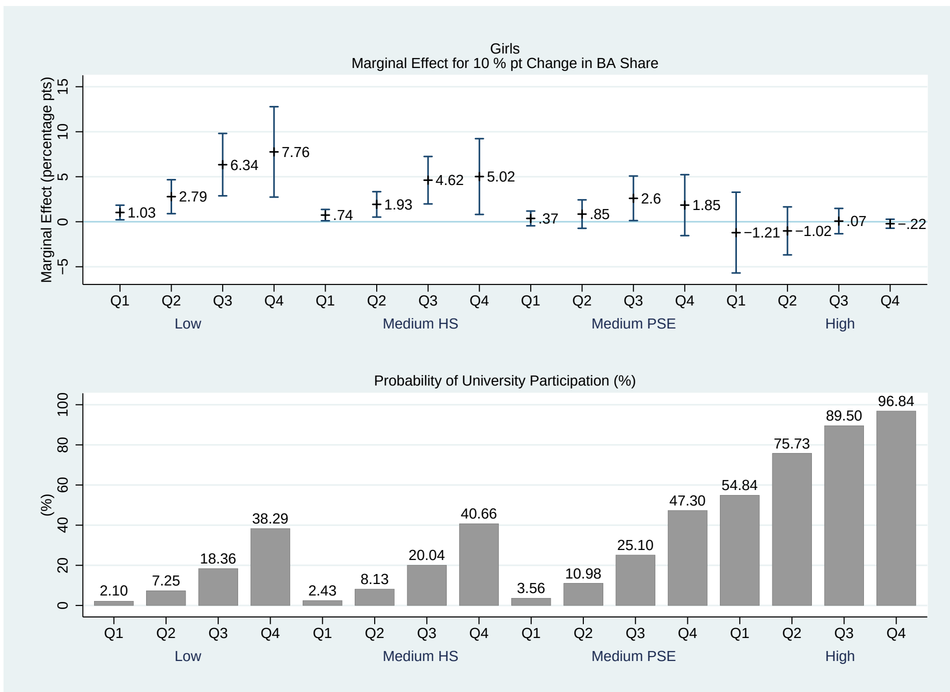

Figure 3.3: Predicted Probability of University Participation and Marginal Neighbourhood Effects by PISA Quartile and SES Type for Girls

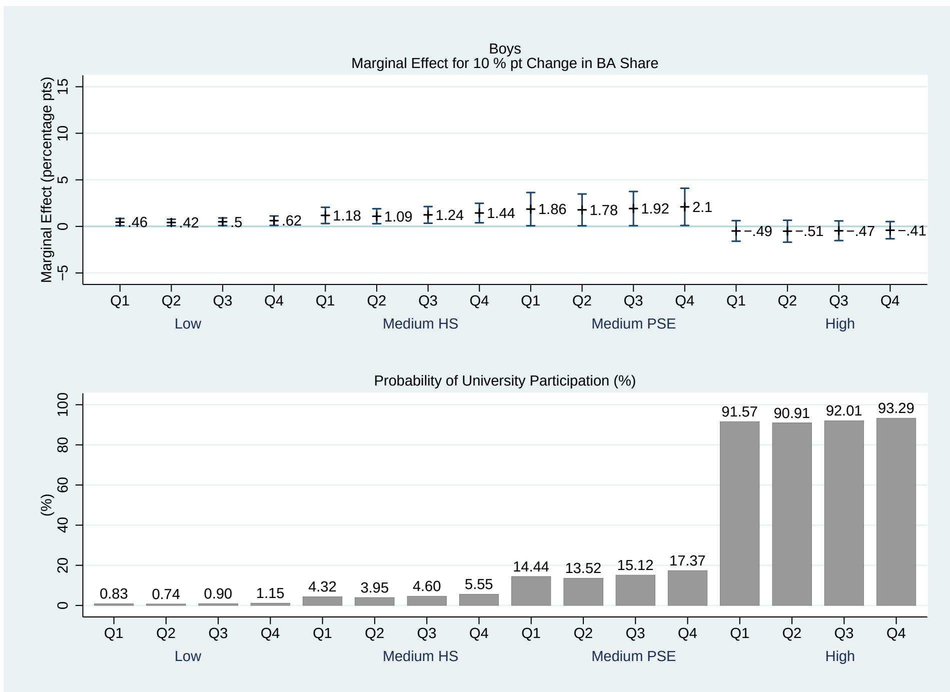

Figure 3.4: Predicted Probability of University Participation and Marginal Neighbourhood Effects by Family Income Quartile and SES Type for Boys

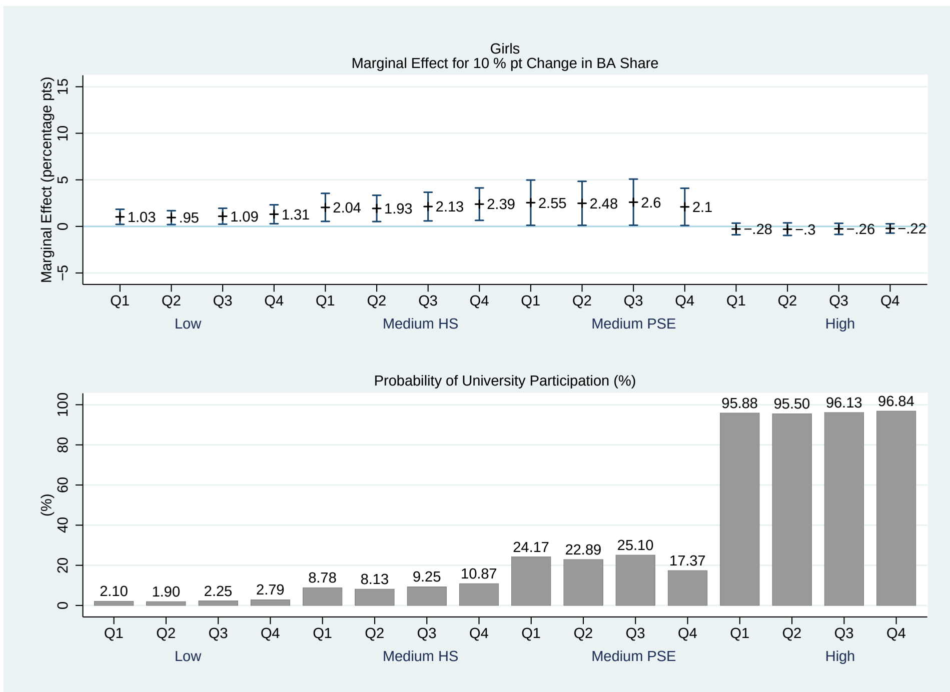

Figure 3.5: Predicted Probability of University Participation and Marginal Neighbourhood Effects by Family Income Quartile and SES Type for Girls

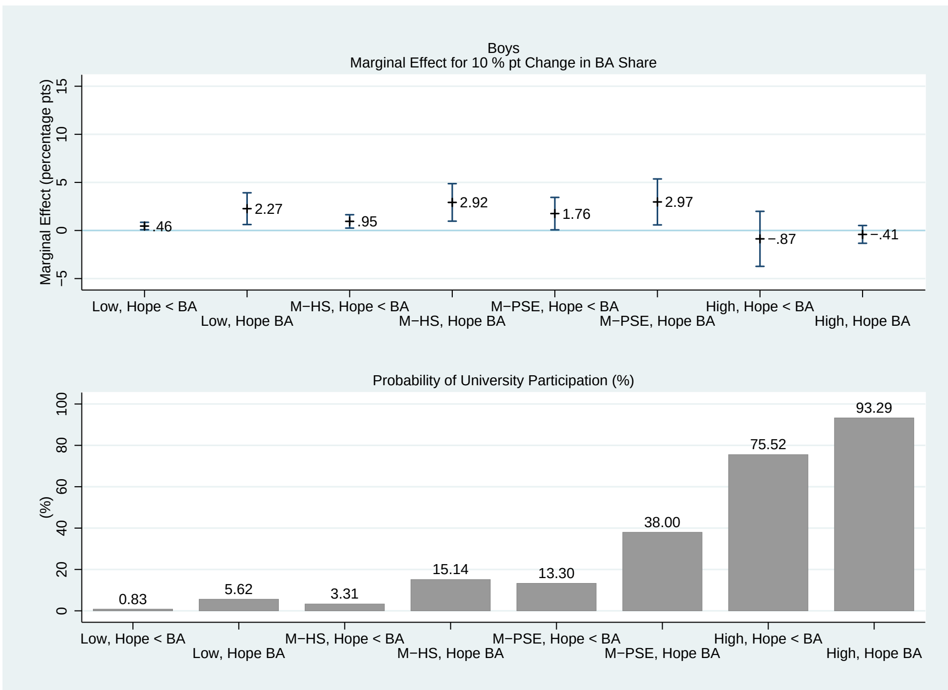

Figure 3.6: Predicted Probability of University Participation and Marginal Neighbourhood Effects by Parents' Hopes for their Children's Educational Attainment and SES Type for Boys

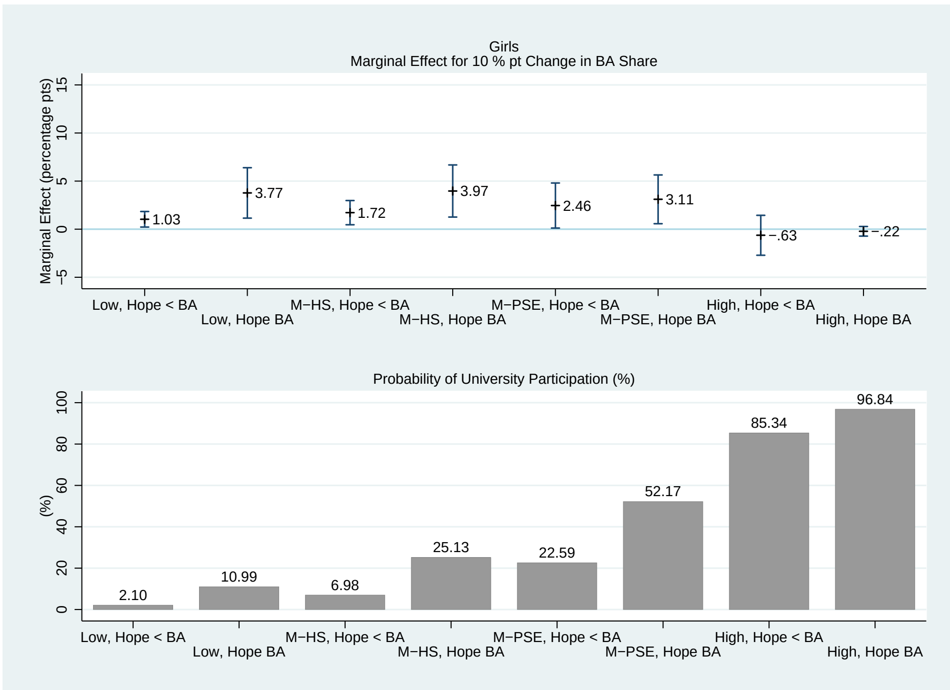

Figure 3.7: Predicted Probability of University Participation and Marginal Neighbourhood Effects by Parents' Hopes for their Children's Educational Attainment and SES Type for Girls

3.6 Bibliography

D. Aaronson. Using sibling data to estimate the impact of neighborhoods on children's educational outcomes. *Journal of Human Resources*, 33(4):915–46, 1998.

G. Akerlof and R. Kranton. Economics and identity. *The Quarterly Journal of Economics*, CXV(3):715–753, 2000.

L. Barrow. School choice through relocation: Evidence from the Washington, D.C. area. *Journal of Public Economics*, 86(2):155–189, 2002.

P. Bayer, S. Ross, and G. Topa. Place of work and place of residence: Informal hiring networks and labor market outcomes. NBER Working Papers W11019, National Bureau of Economic Research, Inc, 2005.

P. Bayer, F. Ferreira, and R. McMillan. A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4): 2001–2052, 2007.

P. Belley and L. Lochner. The changing role of family income and ability in determining educational achievement. *Journal of Human Capital*, 1:37–89, 2007.

B. Bernheim. A theory of conformity. *Journal of Political Economy*, 102(5):841–877, 1994.

A. Bisin and T. Verdier. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97:289–319, 2001.

S. Black, P. Devereux, and K. Salvanes. Why the apple doesnt fall far: Understanding intergenerational transmission of human capital. *The American Economic Review*, 95:437–449, 2005.

A. Boniskowa. Educational attainment of three generations of immigrants in Canada: Initial evidence from the Ethnic Diversity Survey. Discussion Paper 05-21, UBC Department of Economics, 2005.

G. Borjas. Ethnicity, neighbourhoods, and human-capital externalities. *American Economic Review*, 85(3):365–390, 1995.

W. Brock and S. Durlauf. Identification of binary choice models with social interactions. *Journal of Econometrics*, 140:52–75, 2007.

J. Brooks-Gunn, G. Duncan, P. Kato-Klebanov, and N. Sealand. Do neighbourhoods influence child and adolescent development? *The American Journal of Sociology*, 99(2):353–395, 1993.

S. Canada. University tuition fees. *The Daily*, 2005. URL http://www.statcan.ca/Daily/English/050901/d050901a.htm. September 1, 2005.

Canadian Undergraduate Survey Consortium. Survey of first year university students. Technical report, CUSC, Ottawa, 2001.

D. Card and A. Krueger. *Does Money Matter? The Link Between Schools, Student Achievement and Adult Success*, chapter Labor Market Effects of School Quality: Theory and Evidence. Brookings Institution, Washington, DC, 1996.

D. Card and J. Rothstein. Racial segregation and the black-white test score gap. *Journal of Public Economics*, 91:2158–2184, 2007.

P. Carneiro and J. Heckman. The evidence on credit constraints in post-secondary schooling. *Economic Journal*, 482(705-734), 2002.

L. Christofides, J. Cirello, and M. Hoy. Family income and post-secondary education in Canada (1975-1993). *The Canadian Journal of Higher Education*, 31(1): 177208, 2001.

S. Clampet-Lundquist, K. Edin, J. Kling, and G. Duncan. Moving at-risk teenagers out of high-risk neighborhoods: Why girls fare better than boys. Working Paper 509, Princeton IRS, 2006.

J. S. Coleman. Equality of educational opportunity study. Technical report, U.S. Department of Health, Education, and Welfare, Office of Education, 1966.

M. Corak, G. Lipps, and J. Zhao. Family income and participation in post-secondary education. Working Paper 997, IZA, 2004.

M. Corcoran, R. Gordon, D. Laren, and G. Solon. The association between men's economic status and their family and community origins. *Journal of Human Resources*, 27(4):575–601, 1992.

L. Curtis, M. Dooley, and S. Phipps. Child well-being and neighbourhood quality: Evidence from the Canadian National Longitudinal Survey of Children and Youth. *Social Science and Medicine*, 58(19171927), 2004.

D. Datcher. Effects of community and family background on achievement. *The Review of Economics and Statistics*, 64(1):32–41, 1982.

T. Drewes. Returns to college education: Evidence from the 1990, 1995, and 2000 National Graduates Surveys. Technical report, Learning Policy Directorate, Human Resources and Social Development Canada, 2006.

M. Drolet. Participation in post-secondary education in Canada: Has the role of parental income and education changed over the 1990s? Analytical studies research paper series, Statistics Canada, 2005.

G. Duncan, J. Boisjoly, and K. M. Harris. Sibling, peer, neighbor, and schoolmate correlations as indicators of the importance of context for adolescent development. *Demography*, 38(3):427–447, 2001.

G. Duncan, J. Boisjoly, M. Kremer, D. Levy, and J. Eccles. Peer effects in drug use and sex among college students. *Journal of Abnormal Child Psychology*, 33 (3):375–385, 2005.

S. Durlauf. *Handbook of Regional and Urban Economics*, chapter Neighborhood effects, pages 2173–2242. Elsevier, 2004.

A. Ferrer and W. Riddell. The role of credentials in the canadian labour market. *Canadian Journal of Economics*, 35:879–905, November 2002.

R. Finnie, C. Laporte, and E. Lascelles. Family background and access to post-secondary education: What happened over the 1990s? Analytical Studies Branch Research Paper Series 2004226e, Statistics Canada, Analytical Studies Branch, 2004.

M. Frenette. Access to college and university: Does distance matter? Analytical Studies Branch Research Paper Series 11F0019MIE2003201, Statistics Canada, Analytical Studies Branch, 2003.

R. G. Fryer and P. Torelli. An empirical analysis of 'acting white'. NBER Working Papers W11334, National Bureau of Economic Research, Inc, Mar. 2005.

S. Gibbons. Neighbourhood effects on educational achievement, identification using the Census and National Child Development Survey. Discussion Paper LSE 18, Centre for Economics of Education, 2001.

D. Goux and E. Maurin. Close neighbours matter: Neighbourhood effects on early performance at school. *The Economic Journal*, 117:1193–1215, 2007.

C. M. Hoxby. Peer effects in the classroom: Learning from gender and race variation. NBER Working Papers 7867, National Bureau of Economic Research, Inc, 2000.

S. Junor and A. Usher. The price of knowledge 2004: Access and student finance in Canada. Technical report, Millennium Scholarship Foundation, Montreal, 2004.

M. Keane and K. Wolpin. The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review*, 42(4):1051–1103, 2001.

J. Kling, J. Liebman, and L. Katz. Experimental analysis of neighborhood effects. NBER Working Papers W11577, National Bureau of Economic Research, Inc, 2005.

D. E. Kohen, J. Brooks-Gunn, T. Leventhal, and C. Hertzman. Neighbourhood income and physical and social disorder in Canada: Associations with young children's competencies. *Child Development*, 73(6):1844–1860, 2002.

J. Ludwig. Information and inner city educational attainment. *Economics of Education Review*, 18:17–30, 1999.

C. Manski. Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60:531–542, 1993.

K. Milligan. *Higher Education in Canada*, chapter Who uses RESPs and Why, pages 467–494. McGill-Queen's University Press, 2005.

R. Moffitt. *Social Dynamics*, chapter Policy Interventions, Low-Level Equilibria, and Social Interactions. MIT Press, 2001.

National Center for Education Statistics. Digest of education statistics, 2006. Technical Report NCES 2007017, U.S. Department of Education, 2007.

P. Oreopoulos. The long run consequences of growing up in a poor neighborhood. *Quarterly Journal of Economics*, 118(4):1533–1575, 2003.

P. Oreopoulos. Neighbourhood effects in Canada: A critique. University of Toronto, Mimeo., 2008.

P. Oreopoulos, M. Page, and A. H. Stevens. Does human capital transfer from parent to child? The intergenerational effects of compulsory schooling. *Journal of Labor Economics*, 24(4):729–760, October 2006.

L. Orr, J. Feins, R. Jacob, E. Beecroft, L. Sanbonmatsu, L. Katz, J. Liebman, and J. Kling. Moving to opportunity interim impacts on evaluation. Technical report, U.S. Department of Housing and Urban Development, 2003.

J. Ries and T. Somerville. School quality and residential property values: Evidence from vancouver rezoning. Department of economics working paper, University of British Columbia, 2004.

R. Tremblay, B. Boulerice, H. Foster, E. Romano, J. Hagan, and R. Swisher. Multilevel effects on behaviour outcomes in Canadian children. Applied Research Branch of Strategic Policy Working Paper W-01-2E, Human Resources Development Canada, 2001.

S. Venkatesh. *American Project: The Rise and Fall of a Modern Ghetto.* Harvard University Press, 2000a.

S. Venkatesh. *The Robert Taylor Homes Relocation Study*. The Center for Urban Research and Policy, New York, 2000b.

J. Vigdor. *Deviant Peer Influences in Programs for Youth: Problems and Solutions*, chapter Peer Effects in Neighborhoods and Housing. Guilford Press, New York, 2006.

J. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. MIT Press., Cambridge, 2002.

Chapter 4

Can Middle Achieving Students Succeed in Advanced Classes? Impacts on High School Achievement from the BC AVID Pilot Project[1]

4.1 Introduction

The idea that having an educated populace benefits society in terms of economic and social outcomes is fairly uncontroversial. Achieving educational outcomes at the post-secondary level generally requires some academic success in high school courses and examinations. Consequently, some students may be impeded in gaining higher levels of education by their achievement in high school courses. This suggests an avenue for policy makers to increase education levels through improving achievement in high school grades.

There are at least two ways to think about how to improve high school achievement. One way of thinking about the issue would argue that students who strug-

[1] A version of this chapter will be submitted for publication. K. Foley. Can Middle Achieving Students Succeed in Advanced Classes? Impacts on High School Achievement from the BC AVID Pilot Project

gle need to tackle material at their current skill level and move at a slower pace to ensure that they do not lose confidence. One example of this approach is the controversial policy in the Canadian province of Ontario which allows students to submit assignments without penalty when they miss a deadline. Another way of approaching high-school achievement assumes that students should be held to higher standards, and that students will rise to meet those expectations. The 'No Child Left Behind' Act introduced in 2001 in the United States implicitly takes this approach in attempting to close the racial achievement gap by applying the same achievements standards to all students.

The British Columbia Advancement Via Individual Determination (BC AVID) Pilot Project combines both of these approaches to high school achievement. BC AVID is an effort to demonstrate and evaluate a post-secondary preparatory course designed to increase access to post-secondary education for high school students who experience achievement-related barriers. BC AVID encourages students, who upon entering grade 8 are achieving in the middle range, to take advanced classes and offers them the chance to enroll in an elective class that provides academic support, tutoring and information about post-secondary opportunities.

BC AVID is evaluated using a random assignment methodology. Random assignment means that students who volunteer for the program and are found eligible are allocated by a chance lottery into two groups. If assigned to a program group the students are invited to enroll in the AVID elective and participate in all the programs and services associated with the course. Students assigned to the control group are not permitted to enroll in the AVID course but can participate in any of the services or programs normally available to them. Because participation in the program is determined randomly, the program impact can be estimated by comparing the mean outcome within the program group to that within the control group.

In this paper, I examine the impact of BC AVID on achievement in grade 10 courses for which students are required to write provincial examinations. I focus on achievement in math courses because it is possible to distinguish between math courses that are more challenging than others. Specifically, I examine the extent to which students offered a place in the BC AVID elective are likely to take the Principles of Math course, which is the most challenging among the math courses. I also estimate the programs' impact on the chances that students will pass their provincial exam in math, their course work and whether they obtain the final credit.

I find that BC AVID encouraged more students to take the Principles course, but also increased the chances that students would fail the provincial exam in Principles. Students were, however, more likely pass the course work and as a consequence earn credit for Principles of Math. Because Principles of Math is required for most university programs in British Columbia encouraging more students to obtain this high school credit may foreshadow a future impact on university participation. However, the fact that BC AVID also increased the chances that students failed their provincial examination might imply that students were encouraged into a level of math beyond their capacity and will struggle in the future.

I explore these results by examining the pre-program characteristics of students who took Principles of Math because they were offered a place in BC AVID. I identify these marginal students by taking advantage of random assignment into the program. Random assignment ensures that the average characteristics of students in the control group are similar to those in the AVID group. Students in the control group who took Principles of Math did so without being offered a place in BC AVID. Within the AVID group, there is a hypothetical group just like those control group students taking Principles, that is a group of students who would have taken Principles in the absence of BC AVID. Although I can not identify this hypothetical group, random assignment ensures that their mean pre-program characteristics are similar to the students taking Principles in the control group. Among the students who took Principles, if there are differences in the mean characteristics of the AVID group relative to control group students, those differences must come from the characteristics of AVID group students who took Principles because of AVID. I use this idea to compare the characteristics of control group students who took Principles to those marginal AVID students inferred to have taken Principles because of AVID.

When compared to control group members who took the Principles of Math course, I find that the marginal AVID group students have similar average scores on standardized tests taken in Grade 7. The marginal AVID students do differ from control group Principles students in some of their demographic and background characteristics. On average, the marginal AVID student is more likely to be a boy and more likely to come from a lone-parent family. The group of students who I infer took Principles because of AVID are less likely to have university level aspirations at the time of random assignment. This finding suggests that BC AVID might provide encouragement and support to youth who might not normally be interested in pursuing university level education. While the results in this paper present only short-run impacts, the general pattern of impacts suggest that BC AVID has, at least initially, helped students who may have lacked the aspirations for higher levels of education to obtain credit for an advanced math course.

The rest of the paper proceeds by first describing in more detail the BC AVID pilot project, including how schools and students were recruited to participate in the program. The next section describes previous research concerning the determinants of high school achievement and evaluations of programs similar to BC AVID. I then discuss in general some issues related to evaluating programs implemented in a school setting. The section following that describes the estimation method which takes into account the issues raised in the previous section. I then describe the data and explain the various paths of Mathematics available to students in British Columbia. Finally, I present results on achievement in Grade 10 courses in which a provincial examination is required. I focus on achievement in Math, and report some analysis which describes the pre-program characteristics of the students who are inferred to have taken an advanced math course because of BC AVID.

4.2 BC AVID

BC AVID is a post-secondary preparatory class based on the *Advancement Via Individual Determination* (AVID) model which was first pioneered in California in the early 1980s. BC AVID targets students who, before entering high school, are achieving grades in the middle range (B's and C's) but have aspirations to attend post-secondary education (PSE). Students are encouraged to enroll in the most 'advanced' classes in their schools' core curriculum and are academically supported through the AVID elective course. The AVID course is offered in Grade 9 through Grade 12. The class is scheduled within the normal course schedule and ideally meets for one hour daily.[2] Students must give up another elective to enroll

[2]The scheduling systems varied across project sites. Some sites run two separate semesters, while others operate on a 'linear' system where the same classes are taken throughout the academic year. The AVID elective was integrated into these schedules so that at most sites students attended the class at least every other day for the entire year.

in BC AVID.[3]

In the class, a specially trained AVID teacher delivers a curriculum designed to help students acquire the skills and techniques necessary to succeed in more challenging academic classes. The curriculum was developed by the San Diego based AVID Center, which trains teachers, helps initiate programs and provides certification of programs that are found to have adhered with the essentials of the AVID program. The primary components of the curriculum are called "Writing, Inquiry, Collaboration and Reading" (Swanson et al. 2004). Students receive instruction in specific note-taking and critical reading techniques. They are also encouraged to learn by working in groups and engaging in debates.

In addition to the curricular component of the AVID elective, students also receive assistance from tutors, ideally drawn from among local PSE students. Students may also participate in field trips to local businesses or take tours of post-secondary institutions. The AVID model advocates a ratio of 40 per cent curriculum, 40 per cent tutorial and 20 per cent motivational activities.

School Recruitment

Because BC AVID is a demonstration or pilot project, rather than an evaluation of an existing project, evaluators needed to recruit high schools to participate in the project.[4] In 2004, the British Columbia Ministry of Education invited school districts across the province to apply to participate in the BC AVID project. From among 28 applications, 18 high schools in 15 different school districts were selected. Of those 18 sites, four participated in a case study evaluation. The random assignment evaluation was undertaken at 14 high schools and also involved the participation of several middle schools that recruited students in grade 8.[5]

The committee responsible for selecting sites, which included representatives from the Millennium Scholarship Foundation (the funders), the Social Research

[3] The BC curriculum includes a course called Planning 10 in which students learn about various career and post-secondary options. The content of this course was incorporated into AVID. This meant that credit could not be received for both AVID and Planning 10.

[4] When the demonstration project began, AVID was being offered in the Chilliwack school district in BC but the program in Chilliwack was not included in the evaluation.

[5] The grade levels which constitute 'high school' differ across schools in British Columbia. At two participating sites, the middle school fed into the high school in grade 10 while at one site students transferred from middle school to high school in grade 11. In 9 of the 14 sites, students were at the same school from grade 8 to grade 12. More details about the structure of the participating schools are reported in Dunn et al. (2009; Chapter3).

and Demonstration Corporation (the evaluators), the Ministry of Education, and the AVID Center, took into account school size and geography when making their decisions. The intent was to recruit schools with potentially large populations of students suitable for AVID. Because AVID had originally been developed for large urban schools with significant populations of children from lower socio-economic backgrounds, the selection committee had hoped to attract schools with that profile. Unfortunately, the schools that applied tended to be smaller, less urban and the students enrolled came from higher socio-economic backgrounds.

The recruitment of schools gives rise to at least two important implications. First, within the school's population there may not be enough students with the attributes designated as suitable for AVID. This would compromise the possibility of implementing the program as designed. Second, if the impact of AVID differs across schools then the overall estimates of the program's effectiveness will depend on the schools participating in the evaluation. This does not affect the consistency or internal validity of the estimates, but is related to external validity or the extent to which one can extend the estimates to other populations.

Student Recruitment and Selection of BC AVID Eligible Students

The process used to recruit participants for the BC AVID pilot project is of particular importance because BC AVID is not designed as a program for all students. Indeed, the AVID Center considers the selection of 'AVID eligible' students to be a defining program characteristic. One of the program's primary goals is to increase high school achievement for students who are performing below their potential. For some students achieving a grade level of B or C represents an appropriate outcome given their innate ability and talents. Even with the support of the AVID elective, these students may not be able to cope in advanced classes. At the other end of the achievement distribution students who achieve grades in the A range may not need the support provided in BC AVID.

Motivation is another important component in high school achievement. Some students may not exert enough effort to receive the grades that would match their academic potential. For these students their current level of effort might represent the optimal allocation of their time and energy given their future aspirations. Since BC AVID does not seek to change the aspirations of students, the program targets students who would like to pursue PSE and those who have not demonstrated any behavioural issues at school.

Because BC AVID is unlikely to make a positive difference for all students, the effectiveness of the program is vitally linked to identifying students who are motivated but are achieving below their potential. Using broad principles taken from the AVID Center, a committee, including the AVID Center and the evaluators at SRDC, developed selection criteria and procedures for the BC AVID pilot project. Although selection would occur at each site, the procedures were standardized to ensure as much uniformity across sites as possible. Despite these efforts the process of selection may have led to unobservable differences across schools in the composition of those invited to enroll in the AVID elective.

The procedure to identify and recruit suitable students occurred over the year in which students were enrolled in Grade 8. The first step was to inform all grade 8 teachers in the participating schools of the selection criteria and procedures. Teachers were then asked to identify and recommend students who they thought were suitable and to encourage those students to apply for the program. Since all teachers who had contact with grade 8 students were asked to make recommendations, many of these teachers had no previous involvement with the BC AVID pilot project. Despite briefings to inform teachers of the appropriate selection criteria, evidence from depth interviews at the research sites suggests that some teachers were not well informed of the ideal BC AVID student.[6]

Program administrators at some schools also reported that after observing project participants in the first cohort, teachers were better able to identify and recommend suitable students in the second cohort. Quantitatively, there is limited evidence of differences in the observable characteristics of students in the first and second cohorts. Students in cohort one were about six percentage points less likely to have achieved an average grade of B or higher in grade 8, and nearly twice as likely to have received some academic support since the beginning of the grade 8 school year (Dunn et al. 2009). These differences were statistically significant, but no significant differences across the two cohorts were found in the socio-economic background of the students.

All grade 8 students, whether recommended or not, were informed of the application procedures and criteria during school assemblies. Students who self-referred

[6]The evaluation of BC AVID includes the collection of extensive qualitative data documenting the implementation of the program. Dunn et al. (2009) include more information on this data and the methods used for collection.

or were recommended were then invited along with their parents to information sessions where the BC AVID program and evaluation was described in detail. Students were informed at this stage that, if they were selected as eligible to participate in the BC AVID pilot project, they would be randomly allocated into a program group who could enroll in the BC AVID elective or to a control group who would have access to all the other classes and services for which they were normally eligible. At the information session, interested students and their parents were asked to sign an informed-consent form and applications for the program were circulated.

A selection committee at each site, which included the teachers who would deliver the BC AVID elective, reviewed the applications and conducted interviews with the students. The interviews formed an important part of the selection criteria and intended to assess the students' "motivation to achieve personal goals; determination and persistence; commitment to learning and undertaking learning challenges; efficacy, self-awareness and optimism; varied interests; predisposition to learn with others; ability to communicate thoughts and feelings; and support from family" (Dunn et al. 2009; p.45). For consistency across sites, the interviews consisted of 9 scripted questions.

The final selection decision was based on a point system. The highest points were received for falling within the B to C range in grade 8 achievement, having good attendance in grade 8, and students with a B grade or lower were awarded points for having met or exceeded expectations on grade 7 standardized tests. The quality of the interview and written responses in the application were also assessed and awarded points. Students from family backgrounds that might present barriers to participation in post-secondary education also received points, including residing in a single-parent family, in a family with more than 6 members, or in a family with parents who had not participated in PSE. Students identified as Aboriginal or who speak English as a second language also received points for these characteristics.

Students who scored more points than a consistent threshold were deemed eligible for the program and were then randomly assigned. Acceptance rates across the schools and cohorts varied by 72 to 100 percent. Overall 91 per cent of all students who applied were found eligible. Some teachers involved in recruiting students reported that the criteria were too broad and did not effectively filter students. In contrast, other BC AVID teachers felt that they had recruited students who were well suited for the program.

4.3 Previous Research on School Achievement

The BC AVID program implicitly assumes that children's achievement can be affected by the way that children are taught and the content of their curriculum. One branch of the economics literature has thought of this process conceptually as the production of cognitive outputs by combining school and family inputs. In this context, BC AVID can affect children's outcomes by changing the nature of the school inputs. The literature has not always demonstrated a consensus on whether school inputs matter and how they might affect outcomes.

Many reviews of the literature that attempt to explain differences in high school achievement chronologically begin with the *Equality of Educational Opportunity Study* usually referred to as the 'Coleman report' (1966). The Coleman report documented the extent to which schools were segregated in the United States and is often interpreted as suggesting that schools have little effect on the outcomes of children. This interpretation of the report would argue that what determines children's outcomes is their socio-demographic background, not school inputs.

Research relating time variation in expenditures to standardized test scores led Hanushek (1997) to similarly conclude that school inputs were not strong predictors of achievement. In later research, Rivkin, Hanushek, and Kain (2005) found that teachers do matter but the impact is not related to observable characteristics such as credentials but rather because of heterogeneity in unobserved teacher quality. As Todd and Wolpin (2003), point out many of the differences in findings within this literature stem from differences in the theoretical and empirical specifications.

The idea that it is not school inputs but socio-economic background which determines schooling outcomes is particularly relevant in the context of BC AVID. The AVID programs in the United States, which are anecdotally regarded as successful programs, draw almost exclusively from populations of disadvantaged and visible-minority students. While BC AVID recruits students from disadvantaged backgrounds, the program does not place strong emphasis on those characteristics and in practice many participants do come from well-educated, middle to high-income families. The demographic profile of participants might limit the impact the BC AVID has on achievement outcomes.

Because BC AVID does not incorporate any specific variation in different types of school inputs, the evaluation of BC AVID is not particularly suited to contribute to the literature on education production functions. Instead, the evaluation will establish whether one particular combination of inputs affects the achievement of program participants. Previous evaluations of AVID have relied more on qualitative methods and have not identified causal effects as is possible with random assignment evaluation. Dunn et al. (2009) provide a summary of these evaluation efforts. There are, however, random assignment studies of similar programs in the United States. Two programs, Career Academies and Upward Bound, are particularly relevant because of their similarity to AVID in their goals and their participants.

Career Academies are designed to mitigate some of the weaknesses inherent in large comprehensive high schools. Three key principles are common across Career Academies. The first principle refers to a structure described as a 'school within a school'. To provide greater support and engender more engagement, students are organized into classes who remain together and are taught by the same teacher throughout their tenure in high school. The second principle promotes the integration of vocational and academic studies. Finally, Career Academies forge partnerships with local employers to develop career awareness and offer work-based learning opportunities. Unlike BC AVID, in Career Academies there is less emphasis on enrolling in advanced classes and Career Academies target a broad cross-section of students. In practice, however, Career Academies tend to be located in disadvantaged school districts with high concentrations of low-income and African-American students.

As is the case with BC AVID, random assignment in Career Academies was implemented at the level of individual students. At nine Career Academies across the United States, which had received double the number of applications, admission was determined by a random draw.[7] Because participation was determined randomly the impact of the program was estimated consistently by comparing mean outcomes among youth invited to enroll in the Career Academies to the mean outcomes among youth assigned to a control group.

The impact of the program varied substantially with the extent to which youth

[7] There were initially ten sites, however, one site disbanded before the completion of the evaluation.

were characterized as at-risk of high school dropout before random assignment (Kemple 2001). For youth who initially had a high probability of dropping out, by the end of grade 12, Career Academies reduced the chances of dropping out and also improved attendance, and applications to 2 and 4-year colleges. Career Academies did not, however, improve scores on standardized math or reading tests.

In contrast, Career Academies had little effect on the high school outcomes of youth who entered the program with a low to medium chance of dropping out of high school. For these youth, the difference between Career Academies and what was regularly offered at the school may not have been sufficiently large to make a difference for their outcomes. The program consistently reduced the dropout rate in schools where, relative to the control students, Academy students reported high levels of support. In schools, where support reported by Academy students was similar to control group students, the program actually increased high school dropout rates. Looking across all students and all sites, the estimated impact of the program is not statistically different from zero on high school achievement and completion outcomes.

Upward Bound bears some relevant similarities to BC AVID not just in the content and goals of the program but also in some of aspects of the evaluation design. Upward Bound, unlike AVID, does not operate within high schools but is instead hosted at various post-secondary intuitions across the United States. Upward Bound programs aim to increase the skills and motivation of youth who come from low-income families and are potentially first-generation participants in post-secondary education. The programs offer a variety of services including instruction, tutoring and counseling. Participants attend regularly scheduled meetings through the academic year and can participate in a six week intensive instructional program during the summer. Students can participate in Upward Bound from grade 9 through to grade 12, although most participate for less than two years.

The evaluation of Upward Bound was initiated in 1991 and involved a random sample of programs from across the United States (Seftor et al. 2009). Because the evaluation was based on a random assignment design and also drew from a nationally representative sample of programs, the results represent both internally and externally valid estimates. Students were randomly assigned to the program or control group, after they applied to their local Upward Bound program and were deemed eligible by that program.

The sample of 67 programs that participated in the evaluation was highly stratified and over-represented some groups of particular interest to the evaluators. To account for differences in the probability of selection into the sample, each individual in the sample was weighted by the probability that her project was selected and the probability that the sample member was selected to have the opportunity to participate. Because of the sampling scheme, some of the projects were weighted particularly heavily. Specifically, only one project was selected to represent projects that were medium -sized, located in an urban setting, hosted by a four-year public institution and not serving a group of students that is predominantly Asian, Native American or Latino. Because that type of project represented 26 per cent of programs nationwide, the results from this project were given a correspondingly large weight. The evaluators are careful to test how sensitive their results are to the exclusion of this project.

While Upward Bound did not appear to have large effects on many important outcomes, there were several impacts on outcomes for youth who did not expect to obtain a Bachelors degree at the outset of the program (Seftor et al. 2009). For the average participant, however, the program had no impact on high school grades, credits earned or graduation. Upward Bound did not increase overall post-secondary enrollment, nor did it increase applications or receipt of financial aid, in the form of Pell Grants. Upward Bound did increase the fraction of program group members who had completed a post-secondary certificate or license from a vocational school.

Impacts were much larger among youth who entered the program with low educational expectations. These students were more likely to earn credits in Advanced Placement classes. Upward Bound also increased the fraction of these students who enrolled in a 4-year college or university and also increased the number of college credits earned.

4.4 Estimating the Impact of School-Based Programs

While BC AVID is a program designed to affect individual outcomes, it is delivered within a group setting, specifically within a high school classroom. The group setting contrasts with programs that have largely been the focus of evaluation literature in economics and are delivered by a single institution directly to individual participants. A number of different issues arise when evaluating a program delivered in a school setting. One complication arises because some of the factors that affect outcomes are shared by students within a school. For example, students attending the same school in a particular cohort share common teachers, educational resources and the same social environment. These school-cohort level effects must be taken into account both when estimating the program impact and the standard error of that estimate.

Additionally, because BC AVID is necessarily spread across several sites there could be fundamental differences in the way the program was implemented and the characteristics of the participants. BC AVID participants were not randomly selected from a particular population. Instead, participants were recruited and selected by program staff at each of the schools. Although the selection criteria are well-defined and efforts were taken to ensure that the criteria were interpreted and applied consistently across schools, several dimensions of the selection mechanism are subjective. In particular, students were given more points for a teacher recommendation. Students also obtained points contributing toward selection from the subjective assessment of written and verbal responses to questions about their motivation and support within their home environment.

Evidence from interviews with members of selection committees at each school suggests that the characteristics of the participants are likely to differ across the sites (Dunn et al. 2009). There was some ambiguity in what was understood by teachers responsible for recommending students as suitable for AVID. Consequently, there is scope for different interpretations of what constitutes an AVID eligible student. Even some apparently objective measures, such as grade achievement can differ across schools. A grade of 'C-' might reflect achievement of zero to 59 per cent at a school with a policy of not holding students back in grade levels. In contrast, the same grade at a school that allows a student to 'fail' would reflect achievement in a more narrow range (Dunn et al. 2009).

Another reason that the participants might differ across schools is because the underlying student populations vary across schools. The selection criteria are multi-dimensional and it is possible to obtain a qualifying score while fulfilling only some of the selection criteria. Consequently, differences in the participants' characteristics across schools might occur even if there was no subjectivity in the scoring. For example, an aboriginal student, whose parents are high school dropouts and who was strongly recommended by a teacher, had no behavior problems and good attendance would be found AVID eligible, even if she had very low grade achievement and low standardized tests. Another student with similar recommendations from a teacher, also with good behaviour and attendance would be found eligible for BC AVID, if he came from an advantaged family background and scored well on standardized tests and achieved grades in the B range. Schools with very different populations would in all likelihood generate different types of AVID eligible students.

While there are these reasons to expect that the participants in the pilot project will differ across sites and cohorts, in both observable and unobservable ways, students were randomly assigned separately by school and cohort. This means that within a school and cohort the chance of being assigned to the program group is still random. The method used to estimate program impacts will, however, need to account for the difference in participants' characteristics that occur across schools and cohorts.

In addition to differences in the characteristics of participants, the multi-site nature of BC AVID might lead to differences in the nature of the program delivered. As with the selection procedures, consistency in delivery across sites was given great attention by the program designers and those responsible for implementation. For example, a detailed operations manual was created, and all teachers and program staff were trained at the AVID Center in California. Despite these efforts, some differences are virtually unavoidable. For example, insofar as the characteristics of individual teachers matter, the program would differ across sites along this dimension.

Other, perhaps more important, differences in the manner that BC AVID was delivered are also evident. The tutorial component of BC AVID is an area where implementation of BC AVID appears to be most uneven. Most schools reported difficulty recruiting tutors for the BC AVID tutorials (Dunn et al. 2009). According to the AVID Center, the ideal tutor is a student who is currently pursuing post-secondary education. At some of the BC AVID sites senior high schools students were used instead. Other schools relied on adult tutors drawn from among retired teachers, adult volunteers or educational support staff.

Variation across schools is not necessarily a negative outcome. Allowing for flexibility in program delivery might be an effective design element, particularly when implementing a program in a large and diverse jurisdiction like British Columbia.

However, in terms of evaluating the overall impact of the BC AVID pilot project, it is important to be aware that there might be differences in the impact of the program as a result of variations in program delivery. If the program is more or less effective for different types of students then heterogeneity in the program's impact might also arise from differences in participants across sites. Generally, in this context, it is probably not reasonable to assume that BC AVID will have the same impact at each school and within both cohorts.

4.5 Estimating Impacts in BC AVID

Evaluating the impact of a program is greatly simplified when participation is randomly determined. As the previous section discussed, some complications arise when the program is implemented in a school context. Specifically, the estimator should account for potential differences in the participants' characteristics and the impact of the treatment across schools. Additionally, both the point estimates and the standard errors need to take into account school-cohort level effects.

The discussion of how the impacts of BC AVID are estimated begins by defining any outcome of interest (y_{isc}) for a student $i = \{1, \cdots, N_{sc}\}$ attending school s in cohort c,

$$y_{isc} = f\left(\alpha_0, \text{AVID}_{isc}, \theta_{sc}, \varepsilon_{isc}; \delta_{isc}\right) \tag{4.1}$$

where,

$$\text{AVID}_{ics} = \begin{cases} 1 & \text{if assigned to AVID group} \\ 0 & \text{if assigned to control group} \end{cases}$$

The error (ε_{isc}) is i.i.d. where $\mathbb{E}\left[\varepsilon_{isc}\right] = 0$ by assumption. The mean outcome in the population without the program is the random variable, α_0. Any factors that are common within a school and cohort are captured by the unobserved random variable θ_{sc}.

The impact of the program or the treatment effect for each student is δ_{isc}. Policy makers are generally interested in the average treatment effect across some population or sub-group, which might include only those who were 'treated' or those who actually participated in the intervention. I focus on the average treatment effect across all the students who participated in the project. This parameter is denoted as $\mathbb{E}[\delta_{isc}] = \delta$.

I will assume a linear form for (4.1),

$$y_{isc} = \alpha_0 + \delta_{isc}\text{AVID}_{isc} + \theta_{sc} + \varepsilon_{isc} \tag{4.2}$$

I estimate the program impact separately by each school and cohort allowing for a different program impact and variance. The average treatment effect is the weighted average of the estimates within each school and cohort,

$$\begin{aligned} \hat{\delta} &= \sum_{sc} w_{sc}\hat{\delta_{sc}} && (4.3) \\ \mathbb{V}[\hat{\delta}] &= \sum_{sc} w_{sc}^2 \mathbb{V}\left[\hat{\delta_{sc}}\right] && (4.4) \end{aligned}$$

I take advantage of the fact that students were randomly assigned to the AVID program group within each school and cohort to estimate δ_{isc}. Random assignment ensures that,

$$\begin{aligned} \mathbb{E}[\varepsilon_{isc}|\text{AVID}_{isc}, s, c] &= \mathbb{E}[\varepsilon_{isc}] \\ \mathbb{E}[\theta_{sc}|\text{AVID}_{isc}, s, c] &= \theta_{sc} \end{aligned}$$

This means that the treatment effect can be found by differencing the expectation conditional on being assigned to the control group from the expected value of 4.2 conditional on being assigned to the AVID group,

$$\delta_{isc} = \mathbb{E}[y_{isc}|\text{AVID} = 1] - \mathbb{E}[y_{isc}|\text{AVID} = 0] \tag{4.5}$$

The estimate $\hat{\delta}_{isc}$ is obtained using the sample analog of the expectations. In other words, the program effect is just the difference between the average outcomes within the AVID group and control group. The variance within each school and cohort is estimated using the standard ordinary least squares estimator.

The weights used in (4.3) and (4.4) are chosen to put more weight on impacts that are estimated with proportionally larger samples. The estimates are weighted by the fraction of the total sample represented by the sample of that school and cohort, $w_{sc} = N_{sc}/N$. Results are also reported using a simple unweighted average to demonstrate that the impacts are not sensitive to the choice of weights.

The extent to which these weights will generate an externally valid estimate depends on whether the schools participating in the project represent the type of schools that would participate if the program was implemented at scale. The project schools should be thought of as a selected sample because participation depended on the interest of local teachers and administrators. If the program was implemented in such a way that required all schools in British Columbia to deliver AVID then it is unlikely that estimates using these weights would reflect the average effect of the program across the province.

Estimating the impact within each school and cohort takes into account the possibility that participants and treatment effects may differ and removes the effect of common school-cohort level effects. The program impact can be consistently estimated using other approaches, such as weights that account for differences in the random-assignment ratio and school-cohort fixed effects. These estimators are consistent because students were randomly assigned separately by each school and cohort but the variance based on these estimates is not efficient. If weights are used to estimate (4.1) then the school-cohort effect will be reflected in the error term and will generate a correlation across the errors among students attending the same school in the same cohort. The variance of the error in that regression will combine the true error variance plus the variance of the school-cohort effect, $\sigma_\varepsilon + \sigma_{\theta_{sc}}$

This problem could be avoided by including a fixed effect for each school and cohort. If the effect of the treatment is constant across schools and cohorts then the standard errors should be consistently estimated in the model with fixed effects. Instead, if there is treatment heterogeneity, estimating (4.1) with fixed effects and constraining $\delta_{isc} = \delta$ will not provide consistent estimates of the variance. Essentially, constraining the treatment effect to be constant when in truth it varies, will force the error to absorb any differences in treatment effects. In other words, the error will contain $(\delta_{sc} - \delta)\,\mathrm{AVID}_{isc}$. This treatment-heterogeneity component would be common across the error terms for each AVID group members within a school-cohort and would thus generate correlated errors.

As was discussed earlier, there is considerable qualitative evidence that the program was implemented unevenly across the sites. It is also possible to test whether there are any statistically significant differences in the treatment effects across the schools. In several of the outcomes related to high school achievement including enrolling in higher level math classes and provincial exam scores, the program impact did statistically differ across the schools and cohorts. For these reasons, using an estimator that allows for different treatment effects is preferred.

The estimators described in (4.3) and (4.4) are consistent under the assumption that there are no common factors affecting only a subgroup of students within each school and cohort. If one subgroup of students and no others always worked with the same particularly effective tutor then this would be an example of a violation of the assumption. There is no evidence that this particular example was ever the case.

4.6 Data

This paper uses survey and administrative data. Data used to define baseline characteristics come from surveys completed by pilot project participants and their parents prior to random assignment. Students filled out a paper form, while their parents completed a telephone survey. The data was collected during the recruitment and selection process.

Outcomes in academic achievement are measured using administrative data obtained from the BC Ministry of Education. These data include project participants' exam and course scores in all their examinable courses. Specifically, these courses are Mathematics, English and Science. I use only grade 10 courses taken in grade 10. It might also be interesting to study the extent to which students delayed writing their grade 10 exams until grade 11 or 12. However, because at the time this data was available the second cohort was entering grade 11, it would not be possible to study delaying exams consistently for both cohorts.

The administrative data also includes results from Foundations Skills Assessment (FSA) standardized tests which are taken in grades 4 and 7. Students are examined in the domains of reading comprehension, writing, and numeracy. The tests are intended to inform policy makers, teachers and administrators on how well students are achieving the basic skills inherent in the curriculum. FSA tests should not be taken to measure innate ability directly. Certainly FSA scores will partly reflect ability but also any familial and scholastic inputs into cognitive development. Some parents and schools may even pay particular attention to preparing students for FSA exams.

4.6.1 Mathematics in British Columbia

Graduating from high school in British Columbia means that a student has completed the requirements for a *British Columbia Certificate of Graduation* often referred to as a Dogwood diploma. To earn a Dogwood diploma students must earn credits in Mathematics at the grade 11 level. Beginning in grade 10, there are three different paths in mathematics: Principles, Applications and Essentials. Principles of Math emphasizes formalism and symbolic manipulation with a particular focus on preparing students for calculus (British Columbia Ministry of Education 2000). This course is required for entry into most programs at universities throughout British Columbia. The Applications pathway is in contrast geared toward preparing students for 'certificate programs, diploma programs, continuing education programs, trades programs, technical programs'. Finally, Essentials focuses on developing the level of numeracy needed to function in 'daily life, business, industry, and government'. Given these program descriptions, one can identify Principles as being the most challenging.

Earning credits in some courses requires sitting a provincial examination. Examinations for all three levels of mathematics are mandatory at the grade 10 level. The final course grade gives a weight of 20 per cent to the exam score and 80 per cent to the course grade. To pass the course or exam a student must obtain a score above 50 per cent.

4.7 Results

4.7.1 Characteristics of Program and Comparison Group Members at Random Assignment

This section compares the characteristics of participants in the AVID and control groups measured at baseline, or immediately before the program began. The differences in mean characteristics are estimated using the within school-cohort estimator described in the previous section.[8] The section also reports results from tests for statistical significance in the estimated differences.

There are at least two reasons to test for differences in the baseline characteristics of students assigned to the AVID group and those assigned to the control group.

[8] Dunn et al. (2009) reported the means in the AVID group and control group, re-weighted to account for differences in the random-assignment ratios.

First, reporting these characteristics provides some transparency in the evaluation, allowing the reader to develop confidence in the implementation of random assignment. Second, it is important to test for large differences between the research groups which might have occurred because of sampling variation. Differences in the mean baseline characteristics caused by sampling variation do not affect consistency but might warrant the use of regression adjustment to improve the precision of estimates. Baseline characteristics were measured using data from surveys completed by pilot project participants and their parents prior to random assignment.

Table 4.1 reports mean characteristics within the control and AVID groups measured at baseline, or before the student began the program. The measures reported in the table correspond to characteristics which are important components of the selection procedure. For example, several measures of student motivation are included in the table. Other measures, such as age and gender, are basic demographic variables which are typically correlated with important outcomes. The AVID and control group means are reported in the first and second columns respectively. The third column reports the difference in means and the fourth column reports the standard error. Statistical significance is indicated using stars: one star means the difference is significant at a 10 per cent level and two stars indicates significance at the 5 per cent level.

The more outcomes one tests the more likely it becomes that one or more of the outcomes will be statistically significant. It is not surprising therefore that one of the characteristics in Table 4.1 is statistically significant at the 5 per cent level. Students in the AVID group are less likely to hope that they will achieve at least a Bachelors degree. It is common to use regression adjustment to re-balance the groups and improve the precision of the estimates when some differences in baseline characteristics occur. Because the impact estimator already includes a fixed effect for each school-cohort, regression adjustment is not practical in BC AVID.

4.7.2 Grade 10 Achievement

Table 4.2 reports the impact of BC AVID on achievement in grade 10 examinations for courses in which provincial examinations are mandatory. In the first row of each panel, I show the impact of BC AVID on participating in each course. The second and third rows report the impact on taking the exam *and* either passing (reported in the second row) or failing (reported in the third row). Notice that the sum of the second and third rows is equal to the first row.

The first panel presents effects on participation and achievement in Principles of Math. Principles is the most challenging grade 10 math course. Since one of the goals of the AVID program is to encourage students to enroll in more advanced classes. In the control group, 58 per cent of students wrote the Principles exam. BC AVID increased that proportion by roughly 8 percentage points.

BC AVID had no effect on the fraction of students who wrote the Principles examination and passed it. About 47 per cent of the control group wrote and passed Principles, which is just 2 percentage points lower than the fraction of AVID group students who passed the Principles exam. In contrast, BC AVID did significantly increase the fraction of students who failed the exam. Students who were offered a place in the AVID elective were 6 percentage points more likely to have written and failed the Principles of Math examination.

In the second panel of Table 4.2 results are reported for the Essentials of Math course. Students in the BC AVID group were 7 percentage points less likely to take this examination. BC AVID also reduced the fraction of students who passed the Essentials exam by six percentage points. BC AVID had no impact on the chances that students failed the Essentials exam. The program also had no effect on achievement on the Applications of Math examination, which is shown in the third panel of Table 4.2. AVID group students were no more likely to take the Applications exam nor were they more likely to pass or fail the exam when compared to control group.

The 'wrote exam' variables reported in Table 4.2 are constructed in such a way the math categories are mutually exclusive for each student. In other words, students are counted as only having written one exam in grade 10.[9] This means that if, for example, BC AVID increased participation in only one of math pathways, by construction, the impact would be negative in one or both of the other examinations. It would appear from Table 4.2, therefore, that BC AVID encouraged students who would have taken Essentials to enroll instead in Principles and to write that examination.[10]

[9]Some students (less than 5) did write more than one type of Math exam in grade 10. For these students the more advanced exam is counted.

[10]Other combinations of net movement across the three courses could also produce the pattern of impacts reported in Table 4.2. For example, students who would have taken Applications took

The results in the first two panels of Table 4.2 also suggest that by increasing the likelihood of taking Principles of Math instead of Essentials, BC AVID increased the chances that students would fail their grade 10 examination. Because the fraction of students who passed the Essentials exam fell by 6 percentage points and the fraction who failed their Principles exam increased by the same fraction, one might infer that BC AVID encouraged students who would have passed Essentials to take Principles and as a result they failed their exam. The results presented in Table 4.2 are not positive proof of such a conclusion, because other offsetting combinations of impacts could result in the same net effect. For example, it could be that students who were directly induced by BC AVID to take Principles passed the exam, while students who would have taken the course without BC AVID failed the exam as a result of their participation in the BC AVID. However, the former example is the more likely scenario.

The final two panels of Table 4.2 present the impact of BC AVID on achievement in Grade 10 English and Science provincial examinations. Over 90 per cent of grade 10 students in both the AVID and the control group wrote the Science and English exams. It is not surprising then that BC AVID had no impact on the fraction writing these exams. It is still worth considering whether the program had an impact on the probability of passing these exams. Encouraging students to take rigorous courses is a fundamental component of BC AVID. One unintended consequence might be poor performance in core courses like English and Science because students are over-taxed in more challenging courses like Principles of Math. The last two panels of Table 4.2 suggest that this was not the case. Fewer than five per cent of students fail their English exam and being offered a place in the BC AVID elective made no difference in the chances of failing. Failing the Science exam was much more common. About 21 per cent of the control group and 19 per cent of the AVID group failed their exams; the difference of two percentage points is not statistically significant.

Scores on provincial examinations contribute 20 per cent of the final grade which is used to determine whether the student has earned credits toward their high school diploma. Therefore, it is possible to have failed the provincial examination and still earn the course credit. Tables 4.3 and 4.4 show the impacts on the course work component and the final grade, respectively. The effect that BC AVID had on

Principles instead and students who would have taken Essentials took Applications.

course work in Principles of Math differs from its effect on exam results. The AVID group was only 2.4 per cent more likely to have failed the Principles course work and this difference was statistically insignificant. In contrast, BC AVID increased the fraction of students who passed the Principles course work by 5.6 percent which is marginally significant at the 10 percent level.

Because course work is weighted by 80 per cent, as the first panel of Table 4.4 shows, more students earned the Math Principles credit because they were offered BC AVID. The program increased the fraction of students who earned a Principles credit by 6 percentage points, even though students were more likely to fail the BC AVID provincial exam. One possible explanation for this result is that the support BC AVID provides was sufficient to help students cope with the more challenging course work associated with Principles but did not adequately prepare them for the provincial examination.

The impacts of BC AVID on course work and final grades in Essentials of Math, shown in the second panel of Tables 4.3 and 4.4 are very similar to the impacts on examination results. If AVID was generally able to improve the performance of students in their course work, then one might expect a reduction in the fraction of students receiving a failing grade on their course work in Essentials. However, since failing the course work associated with Essentials of Math is rare within the control group, there is little scope for BC AVID to have had an impact on this margin.

Although there was no impact on the chances that students take Applications of Math, the third panel of Table 4.3 suggests that BC AVID increased the fraction of students who failed and reduced the fraction that passed the Applications of Math course work. The size of this impact on failing Applications course work is very small despite its statistical significance. The pattern of results in the third panel of Table 4.3 suggests that some of the students who would have passed Applications of Math took Principles or Essentials instead.

A student passes a course or exam when she achieves a score above 50 per cent. Considering only the impact that BC AVID had on this margin might mask other impacts at the higher end of the achievement distribution. Results shown in the top three panels of Table 4.5 indicate whether BC AVID increased the chances that students earned either an A or a B in each of the Math pathways for exams, course work and final grades.[11] There is no evidence in Table 4.5 that BC AVID improved the chances that students would achieve an A or B in any of the Math courses. The negative impact on the fraction of students earning an A or B in Essentials of Math reflects the fact that BC AVID decreased the number of students who took that course. Although one can not be certain using only these results, this evidence does suggest that the students who were encouraged to take BC AVID would have been among the high achieving Essentials students had they not been offered a place in the AVID elective.

The pattern of impacts reported in Tables 4.2 to 4.4 give the general impression that BC AVID encouraged more students to take Principles of Mathematics. The results suggest that BC AVID drew students from Essentials of Math into Principles. While fewer students passed the examination, BC AVID did help more students pass the course work and as a result more students obtained credit for Principles of Math in grade 10. BC AVID did not help student achieve better grades in English or Science, but neither did the program increase the chances that students failed those exams and courses.

The existence of a control group who are unaffected by the program is one of the fundamental assumptions underpinning the use of random assignment in program evaluation. In the context of BC AVID, there are several ways in which this assumption might be violated. The control group students might indirectly receive benefits related to the program. For example, the AVID teachers receive special training as a part of the program. These teachers also instruct regular courses in which control group members might be enrolled. Control members might be affected by these teachers if for example, the teachers are more motivated or implement techniques learned during the AVID training.

Control group students might also be indirectly affected by the behaviour of AVID students. This possibility is particularly revelent because it appears that BC AVID changed the composition of the math classes. Changes in the composition of the math classes might affect control group students if the shift across classes caused a change in the average 'quality' of students within each of the math classes. For example, one could imagine that a group of students might switch from Essentials to Principles lowering the average ability in both Essentials and Principles. If

[11] In British Columbia, an A is achieved for a grade of 86% or above and B is for grades between 73 and 85%.

teachers grade course work comparatively then the course grades of control group students would rise.

If 'spill-overs' such as these are substantial then the results reported here are not consistent estimates of the program's impact. It is not even possible to think of the estimates as lower bounds since one can not be sure of the direction of the effect on control group members. When thinking about direct spill-overs, it might make sense to think about the estimated impacts as the effects of the program elements which are excludable. For example, access to tutors and the AVID curriculum. In the context of indirect spill-overs, one might place greater emphasis on outcomes less likely to be affected by these indirect effects, such as provincial exam results, which are graded at a provincial standard by teachers who are unlikely to know the students.

4.7.3 School-Cohort Heterogeneity

The results in Tables 4.2 to 4.4 are estimated assuming that the effect of the treatment varies across schools and cohorts. This section reports impacts separately by school and cohort and investigates how sensitive the results are to the choice of weights. Table 4.6 shows the impact on the probability that students took the Principles of Math course for each school and cohort. To protect the privacy of program participants the schools are not identified. It is evident from the table that there is considerable variation in the program impacts across the schools and cohorts. Because of the small sample sizes, however, many of the impacts are imprecisely estimated. Because of this considerable noise it is difficult to interpret the variation in impacts. For this outcome, the null hypothesis that the impacts are the same across schools and cohorts is rejected only at the 10 per cent level (p-value = 0.0813).

In Table 4.7, I show the impacts on the probability of failing the Principles of Math exam separately by school and cohort. Once again there is substantial variation in impacts across the schools. In this case, it is possible to reject the null hypothesis that impacts do not vary across the schools and cohorts with a higher level of confidence of 1 per cent (p-value= 0.0085).

Because of the substantial variation in program impacts across the schools, it is clear that it would be possible to manipulate the estimate of the average program impact by setting the weights to emphasize a school with a particularly large positive impact. At the extreme, one could give all the weight to a single school or cohort. While that would clearly constitute an unreasonable weight, there are other choices of weights that are arguably sensible. It is important therefore to demonstrate that the results presented here are not particularly sensitive to the choice of weights. An obvious alternative choice of weight is one which gives equal weight to each school and cohort. Tables 4.8 to 4.10 repeat Tables 4.2 to 4.4 using a simple average instead of a weighted average of the school-cohort effects. In this case, the estimated effects are not dependent on the weight chosen and the impacts are quantitatively very similar in both sets of tables.

4.7.4 Characteristics of Students Encouraged to Take Principles of Math

It is not clear whether policy makers should regard these short-run impacts of BC AVID as foreshadowing outcomes that are ultimately desirable. On the one hand, BC AVID increased the fraction students who failed their Principles of Math exam and reduced the number of students who achieved an A or B in Essentials of Math. Taken together these two impacts suggests that BC AVID might have encouraged students into a level of Math beyond their capacity. If these students have relatively low underlying ability, then they might find that they later experience lower post-secondary achievement or lower wages because they did not pursue an educational path which matched their skills. On the other hand, BC AVID increased the fraction of students who passed the course work and obtained credit for Principles of Math. These students may be able to improve their skills over time if they are able to continue taking math at the Principles level. If this were the case, then the increase in the fraction of students who failed the Principles exam may be a temporary impact which would disappear if students improve their skills and meet the standards expected in the Principles of Math path.

Learning more about the students who were encouraged by BC AVID to take Principles of Math can provide some information which is helpful in interpreting what the short-run impacts imply. Unfortunately, it is not possible to observe which individuals took Principles as a direct result of being offered BC AVID. However, I can exploit the fact that because of randomization, on average, both research groups are similar before the program to examine the characteristics of students who were induced to take Principles because of BC AVID.[12]

The students in the control group who took Principles did so without the aid of the BC AVID program. Although I can not identify them individually, I assume that there is a similar subgroup within the AVID group. In other words, among the AVID students who took Principles of Math, there is a group who would have taken the program even if not offered a place in the BC AVID elective. The remaining students took Principles as a direct result of being offered a place in BC AVID. The purpose of this analysis is to learn more about this group of students. Because students were randomly assigned to the AVID and control groups, on average, their characteristics were similar at baseline or before the program began. It follows that the mean characteristics among students who took Math principles course in the control group are similar to that within the hypothetical subgroup within the AVID group who would have taken Principles in the absence of BC AVID. Any differences in the mean characteristics of the AVID Principles students and the control Principles students must come from the characteristics of the marginal students who took Principles because of BC AVID.

It is possible to decompose the average characteristic of AVID students who took Math principles as,

$$\bar{x}^P_{AVID} = p^P \bar{x}^P_C + \left(1 - p^P\right) \bar{x}^P_M \tag{4.6}$$

where the mean characteristics of the AVID and control group members who took Principles are $\bar{x}^P_{AVID}$ and $\bar{x}^P_C$ and $\bar{x}^P_M$ is mean characteristics of the marginal group or the group who took Principles because they were offered BC AVID. The ratio of the number of students in the control group who took Principles to the number of students taking Principles in the AVID group is p^P. Rearranging 4.6 yields an expression for the average characteristics of the marginal group,

$$\bar{x}^P_M = \frac{\bar{x}^P_{AVID} - p^P \bar{x}^P_C}{(1 - p^P)} \tag{4.7}$$

The means estimated in 4.7 are estimated using weights that account for the different random assignment probabilities across each school and cohort.

Table 4.11 reports the average characteristics of the students who took Principles of Math in the AVID and control groups. The third column reports the mean

[12]Michalopoulos et al. (2005) use a similar approach in evaluating a financial incentive program.

characteristics within the inferred marginal group. Along many dimensions the marginal group is very similar to the control group. Notably, the average score on grade 7 Foundation Skills Assessment tests are very similar in the three domains, numeracy, reading and writing. Because it appears that BC AVID did not encourage students with lower skills in grade 7 to take Math Principles, this result suggests that it might be possible for students who failed the Principles exam to 'catch-up'.

If the marginal students do not, on average, have lower skills then it is likely that they face some other barrier which would have discouraged them from taking Principles in the absence of BC AVID. There is evidence in Table 4.11 that the inferred marginal students differ from control group members who took Principles in their demographics and their educational aspirations.

The average member of the marginal group is 6 percentage points more likely to be male relative to control group Principles students. Because boys generally have lower achievement levels in high school (Frenette and Zeman 2007), the larger fraction of male students in the marginal group may help explain why BC AVID increased the fraction of students who failed the Principles exam. Students in the marginal group are also more likely to live in a lone-parent family. Lone parents may have less time relative to two parent families to spend helping their children with home work. Having access to the additional support in BC AVID may help compensate for this deficit of family time.

The students who are inferred to have taken Principles because of the BC AVID offer are less likely than control group Principles students to have university level aspirations. Among students who took Principles of Math in the control group, 71 per cent hoped they would achieve at least a Bachelors degree. In contrast, within the marginal group only 49 per cent had similar educational aspirations at the time of random assignment. Because Principles of Math is required for entrance into university programs in British Columbia, the fact that BC AVID encourages students to take Principles who did not have university level aspirations before being offered a place in BC AVID could foreshadow a change in their aspirations and potentially a change in their eventual participation in university.

The impact of the program on longer run educational and labour market outcomes will depend greatly on whether BC AVID can help students improve their skills over time. While the short-run findings suggest that students are able use the additional support provided by BC AVID to tackle the course work in Principles of Math, BC AVID also increased the fraction of students who failed their exam. In contrast to course work, where students have access to tutors and classmates, the provincial exam environment is one in which students must rely solely on their own skills.

4.8 Conclusion

This paper has explored the impact of BC AVID on achievement in grade 10 classes. One of AVID's goals is to increase post-secondary participation among students with grades in the middle range during secondary school. With that goal in mind, AVID encourages students to take more advanced classes and provides academic support through an elective course. The results presented in this paper suggest that after two years students who were randomly selected to participate in BC AVID were more likely to enrol in an advanced math course.

The students offered a place in the AVID elective were more likely to take Principles of Math in grade 10, but were also more likely to have failed the provincial examination. Although they were less likely to be successful in the examination, BC AVID increased the chances that students passed their course work and as a result more students obtained credit for Principles. Non-experimental analysis suggests that, relative to Principles students in the control group, the students I infer took Principles of Math as a result of AVID had similar scores on standardized tests taken in Grade 7. However, these marginal AVID students had lower educational aspirations relative to control group students who took principles.

These results provide a snap shot in short-run of how the program is impacting outcomes after two years of participation. It would be premature to draw any concrete policy implications from these results. On the one hand, the program may have unintended impacts in the long run if students are being pushed into courses that do not suit their abilities as the impact on failing provincial examinations might foretell. On the other hand, because this level of Math is required for entry into most university programs in British Columbia, the program's impact on obtaining the credit may constitute an important first step toward other beneficial outcomes. What is clear from the evidence presented in this paper is that BC AVID is having a substantial impact on participants' behaviour and educational outcomes after only two years.

Table 4.1: Differences in Group Characteristics at Random Assignment

	AVID Mean	Control Mean	Difference	Standard Error
Proportion (unless otherwise noted)				
Male (%)	0.477	0.442	0.035	0.030
Age (mean in years)	13.334	13.339	-0.004	0.029
Lone parent	0.206	0.184	0.022	0.024
Aspires to have university degree	0.559	0.619	-0.060	0.029**
Aboriginal	0.089	0.101	-0.012	0.017
English as a second language	0.111	0.111	0.000	0.017
Family income ($)	67,707	68,807	-1,100	2,459
Average grade in B-C range	0.826	0.831	-0.005	0.021
Absent more than 7 days	0.245	0.252	-0.008	0.025
Respondent reports statement is true often or all the time				
I complete my homework on time.	0.720	0.753	-0.033	0.026
I do as little work as possible	0.077	0.079	-0.003	0.016
Sample size	757	432		

Source: Baseline survey data

Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10

Table 4.2: Impacts on Grade 10 Provincial Examinations

	AVID Mean	Control Mean	Difference	Standard Error
Proportion				
Principles of Math				
Wrote Exam	0.661	0.581	0.079	0.028***
Passed Exam	0.492	0.473	0.020	0.030
Failed Exam	0.168	0.109	0.060	0.021***
Essentials of Math				
Wrote Exam	0.057	0.128	-0.071	0.016***
Passed Exam	0.052	0.112	-0.060	0.016***
Failed Exam	0.005	0.016	-0.011	0.006*
Applications of Math				
Wrote Exam	0.152	0.173	-0.021	0.020
Passed Exam	0.128	0.145	-0.017	0.019
Failed Exam	0.024	0.026	-0.002	0.009
English				
Wrote Exam	0.953	0.952	0.001	0.013
Passed Exam	0.908	0.908	0.000	0.017
Failed Exam	0.045	0.041	0.003	0.012
Science				
Wrote Exam	0.928	0.922	0.007	0.016
Passed Exam	0.742	0.716	0.026	0.025
Failed Exam	0.186	0.206	-0.019	0.023
Sample size	757	432		

Source: BC Ministry of Education administrative records
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10

Table 4.3: Impacts on Course Grades for Grade 10 Courses

	AVID Mean	Control Mean	Difference	Standard Error
Proportion				
Principles of Math				
Took course	0.661	0.581	0.079	0.028***
Passed course work	0.590	0.534	0.056	0.029*
Failed course work	0.071	0.047	0.024	0.014
Essentials of Math				
Took course	0.057	0.128	-0.071	0.016***
Passed course work	0.052	0.124	-0.072	0.016***
Failed course work	0.006	0.004	0.001	0.004
Applications of Math				
Took course	0.152	0.173	-0.021	0.020
Passed course work	0.125	0.164	-0.038	0.019**
Failed course work	0.027	0.009	0.018	0.008**
English				
Took course	0.953	0.952	0.001	0.013
Passed course work	0.913	0.898	0.015	0.017
Failed course work	0.040	0.054	-0.014	0.012
Science				
Took course	0.928	0.922	0.007	0.016
Passed course work	0.888	0.878	0.010	0.019
Failed course work	0.041	0.044	-0.003	0.012
Sample size	757	432		

Source: BC Ministry of Education administrative records
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10

Table 4.4: Impacts on Final Grades for Grade 10 Courses

	AVID Mean	Control Mean	Difference	Standard Error
Proportion				
Principles of Math				
Took course	0.661	0.581	0.079	0.028***
Earned credit	0.586	0.527	0.059	0.029**
Failed credit	0.074	0.054	0.020	0.015
Essentials of Math				
Took course	0.057	0.128	-0.071	0.016***
Earned credit	0.053	0.124	-0.070	0.016***
Failed credit	0.004	0.004	0.000	0.004
Applications of Math				
Took course	0.152	0.173	-0.021	0.020
Earned credit	0.128	0.162	-0.034	0.019*
Failed credit	0.024	0.011	0.013	0.008
English				
Took course	0.953	0.952	0.001	0.013
Earned credit	0.923	0.919	0.005	0.016
Failed credit	0.029	0.033	-0.004	0.010
Science				
Took course	0.928	0.922	0.007	0.016
Earned credit	0.879	0.873	0.006	0.020
Failed credit	0.049	0.048	0.001	0.013
Sample size	757	432		

Source: BC Ministry of Education administrative records
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10

Table 4.5: Impacts on Achievement at the Mean and Top of the Grades Distribution

	AVID Mean	Control Mean	Difference	Standard Error
Proportion				
Principles of Math				
Exam grade A or B	0.066	0.064	0.002	0.014
Course work grade A or B	0.202	0.186	0.016	0.024
Final grade A or B	0.167	0.148	0.018	0.022
Essentials of Math				
Exam grade A or B	0.007	0.016	-0.010	0.006
Course work grade A or B	0.012	0.039	-0.028	0.009***
Final grade A or B	0.005	0.033	-0.028	0.007***
Applications of Math				
Exam grade A or B	0.024	0.013	0.011	0.008
Course work grade A or B	0.039	0.042	-0.002	0.011
Final grade A or B	0.030	0.031	-0.001	0.010
Scores in any Math course (%)				
Mean exam score	57.775	59.011	-1.236	0.800
Mean course work score	64.305	65.349	-1.044	0.888
Mean Final score	63.021	64.108	-1.086	0.813
Sample size	757	432		

Source: BC Ministry of Education administrative records
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10

Table 4.6: Impacts on Taking Principles of Math By School and Cohort

	AVID Mean	Control Mean	Difference	Standard Error	AVID N	Control N
School-Cohort						
1	0.632	0.500	0.132	0.107	57	34
2	0.593	0.650	-0.057	0.146	27	20
3	0.533	0.389	0.144	0.151	30	18
4	1.000	0.385	0.615	0.102***	24	13
5	0.414	0.375	0.039	0.156	29	16
6	0.357	0.385	-0.027	0.166	28	13
7	0.556	0.364	0.192	0.144	27	22
8	0.724	0.647	0.077	0.143	29	17
9	0.607	0.619	-0.012	0.144	28	21
10	0.826	0.818	0.008	0.144	23	11
11	0.750	0.294	0.456	0.139***	28	17
12	0.962	0.643	0.319	0.110***	26	14
13	0.862	0.842	0.020	0.106	29	19
14	0.722	0.750	-0.028	0.196	18	8
15	0.690	0.773	-0.083	0.128	29	22
16	0.720	0.636	0.084	0.171	25	11
17	0.407	0.563	-0.155	0.159	27	16
18	0.483	0.571	-0.089	0.166	29	14
19	0.586	0.733	-0.147	0.155	29	15
20	0.667	0.571	0.095	0.166	24	14
21	0.700	0.600	0.100	0.152	30	15
22	0.800	0.733	0.067	0.134	30	15
23	0.808	1.000	-0.192	0.117	26	12
24	0.700	0.533	0.167	0.153	30	15
25	0.680	0.462	0.218	0.168	25	13
26	0.759	0.733	0.025	0.141	29	15
27	0.429	0.250	0.179	0.177	21	12
Sample size	757	432				

Source: BC Ministry of Education administrative records
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10
Schools and cohorts can not be identified specifically to protect the privacy of participants

Table 4.7: Impacts on Failing Principles of Math By School and Cohort

	AVID Mean	Control Mean	Difference	Standard Error	AVID N	Control N
School-Cohort						
1	0.088	0.059	0.029	0.058	57	34
2	0.000	0.150	-0.150	0.070**	27	20
3	0.067	0.000	0.067	0.060	30	18
4	0.208	0.077	0.131	0.129	24	13
5	0.103	0.000	0.103	0.078	29	16
6	0.071	0.077	-0.005	0.090	28	13
7	0.185	0.136	0.049	0.108	27	22
8	0.069	0.353	-0.284	0.110**	29	17
9	0.286	0.190	0.095	0.126	28	21
10	0.304	0.000	0.304	0.143**	23	11
11	0.214	0.059	0.155	0.112	28	17
12	0.538	0.143	0.396	0.154**	26	14
13	0.103	0.105	-0.002	0.092	29	19
14	0.111	0.250	-0.139	0.157	18	8
15	0.241	0.091	0.150	0.108	29	22
16	0.200	0.182	0.018	0.147	25	11
17	0.037	0.063	-0.025	0.068	27	16
18	0.241	0.000	0.241	0.117**	29	14
19	0.103	0.333	-0.230	0.120*	29	15
20	0.125	0.000	0.125	0.091	24	14
21	0.167	0.133	0.033	0.117	30	15
22	0.300	0.200	0.100	0.142	30	15
23	0.192	0.250	-0.058	0.146	26	12
24	0.167	0.000	0.167	0.098*	30	15
25	0.160	0.077	0.083	0.118	25	13
26	0.207	0.000	0.207	0.107*	29	15
27	0.190	0.083	0.107	0.132	21	12
Sample size	757	432				

Source: BC Ministry of Education administrative records
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10
Schools and cohorts can not be identified specifically to protect the privacy of participants

Table 4.8: Impacts on Grade 10 Provincial Examinations-Using Simple Average Weight

	AVID Mean	Control Mean	Difference	Standard Error
Proportion				
Principles of Math				
Wrote Exam	0.665	0.586	0.079	0.028***
Passed Exam	0.492	0.474	0.018	0.030
Failed Exam	0.173	0.112	0.062	0.022***
Essentials of Math				
Wrote Exam	0.058	0.123	-0.065	0.017***
Passed Exam	0.053	0.106	-0.053	0.016***
Failed Exam	0.005	0.018	-0.012	0.006**
Applications of Math				
Wrote Exam	0.147	0.172	-0.025	0.020
Passed Exam	0.123	0.146	-0.023	0.019
Failed Exam	0.024	0.025	-0.001	0.009
English				
Wrote Exam	0.955	0.953	0.002	0.013
Passed Exam	0.908	0.909	-0.001	0.018
Failed Exam	0.046	0.041	0.006	0.013
Science				
Wrote Exam	0.928	0.923	0.005	0.016
Passed Exam	0.750	0.722	0.028	0.025
Failed Exam	0.178	0.201	-0.023	0.022
Sample size	757	432		

Source: BC Ministry of Education administrative records
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10

Table 4.9: Impacts on Course Grades for Grade 10 Courses-Using Simple Average Weight

	AVID Mean	Control Mean	Difference	Standard Error
Proportion				
Principles of Math				
Took course	0.665	0.586	0.079	0.028***
Passed course work	0.594	0.538	0.056	0.030*
Failed course work	0.071	0.048	0.023	0.015
Essentials of Math				
Took course	0.058	0.123	-0.065	0.017***
Passed course work	0.054	0.119	-0.065	0.016***
Failed course work	0.004	0.004	0.000	0.004
Applications of Math				
Took course	0.147	0.172	-0.025	0.020
Passed course work	0.121	0.164	-0.042	0.019**
Failed course work	0.025	0.008	0.017	0.008**
English				
Took course	0.955	0.953	0.002	0.013
Passed course work	0.917	0.897	0.020	0.017
Failed course work	0.038	0.056	-0.018	0.013
Science				
Took course	0.928	0.923	0.005	0.016
Passed course work	0.886	0.876	0.010	0.020
Failed course work	0.043	0.047	-0.004	0.013
Sample size	757	432		

Source: BC Ministry of Education administrative records
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10

Table 4.10: Impacts on Final Grades for Grade 10 Courses-Using Simple Average Weight

	AVID Mean	Control Mean	Difference	Standard Error
Proportion				
Principles of Math				
Took course	0.665	0.586	0.079	0.028***
Earned credit	0.590	0.531	0.059	0.030**
Failed credit	0.075	0.055	0.021	0.016
Essentials of Math				
Took course	0.058	0.123	-0.065	0.017***
Earned credit	0.055	0.119	-0.064	0.016***
Failed credit	0.003	0.004	-0.001	0.003
Applications of Math				
Took course	0.147	0.172	-0.025	0.020
Earned credit	0.124	0.163	-0.039	0.019**
Failed credit	0.023	0.009	0.014	0.008*
English				
Took course	0.955	0.953	0.002	0.013
Earned credit	0.926	0.919	0.007	0.016
Failed credit	0.029	0.034	-0.005	0.011
Science				
Took course	0.928	0.923	0.005	0.016
Earned credit	0.878	0.873	0.005	0.020
Failed credit	0.050	0.050	0.000	0.014
Sample size	757	432		

Source: BC Ministry of Education administrative records
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10

Table 4.11: Characteristics of Students Taking Principles of Math

	AVID Principles Students	Control Principles Students	Inferred Marginal Principles Students
Background Characteristics			
Male	0.492	0.460	0.524
	(0.022)	(0.022)	(0.032)
English as a second language	0.134	0.137	0.131
	(0.015)	(0.015)	(0.021)
Aboriginal	0.072	0.069	0.076
	(0.012)	(0.011)	(0.017)
Lone parent	0.179	0.136	0.222
	(0.017)	(0.015)	(0.025)
Parents do not have PSE	0.340	0.340	0.341
	(0.021)	(0.021)	(0.030)
Aspires to have university degree	0.600	0.713	0.487
	(0.022)	(0.020)	(0.032)
Grade 7 Foundation Skills Assessment Scores (%)			
Numeracy score	54.584	55.979	53.257
	(0.815)	(0.775)	(1.156)
Reading score	67.923	67.971	67.878
	(0.660)	(0.693)	(0.908)
Writing score	51.071	52.030	50.157
	(0.480)	(0.474)	(0.674)
Sample size	250	501	

Means are weighted to account for differences in the random assignment ratio across schools and cohorts.
Standard errors are reported in parenthesis
Source: Baseline survey data and BC Ministry of Education administrative records
Statistical significance indicated with stars: *** .01 level, ** .05 level, * .10

4.9 Bibliography

British Columbia Ministry of Education. The Three Pathways. Website, 2000. http://www.bced.gov.bc.ca/irp/math1012/mapath.htm.

J. S. Coleman. Equality of educational opportunity study. Technical report, U.S. Department of Health, Education, and Welfare, Office of Education, 1966.

E. Dunn, R. Ford, I. Kwakye, J. Hutchison, S. Hebert, K. Foley, and L. Wilson. Bc avid pilot project: Early implementation report. Technical report, Millennium Scholarship Foundation, 2009.

M. Frenette and K. Zeman. Why are most university students women? Evidence based on academic performance, study habits and parental influences. Analytical Studies Branch Research Paper Series. Catalogue No. 11F0019MIE-No. 303, Statistics Canada, 2007.

E. A. Hanushek. Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, 19(2):141–164, 1997.

J. J. Kemple. Career academies impacts on students initial transitions to post-secondary education and employment. Technical report, MDRC, 2001.

C. Michalopoulos, C. Miller, P. Robins, P. Morris, D. Gyarmati, C. Redcross, K. Foley, and R. Ford. Making work pay: Final report on the self-sufficiency project for long-term welfare recipients. Technical report, Social Research and Demonstration Corperation, 2005.

S. G. Rivkin, E. A. Hanushek, and J. F. Kain. Teachers, schools and academic achievement. *Econometrica*, 73(2):417–458, March 2005.

N. Seftor, A. Mamun, and A. Schirm. The impacts of regular upward bound on postsecondary outcomes 7-9 years after scheduled high school graduation. Technical report, Mathematica Policy Research, 2009.

M. Swanson, M. Contreras, D. Cota, and R. Gira. *Implementing and Managing the High School AVID Program: The AVID Teacher/Coordinator Guide.* AVID Academic Press, 2004.

P. E. Todd and K. I. Wolpin. On the specification and estimation of the production function for cognitive achievement. *Economic Journal*, 113(485):F3–F33, February 2003.

Chapter 5

Conclusion

The papers presented in my dissertation focus on three distinct aspects of the factors and processes that determine educational attainment among Canadian youth. The first paper emphasized the importance of ability and the parental valuation of education in influencing children's decisions to drop out of high school. The next paper focused on the potential impact of peers and role models, discovering that for highly-skilled youth, whose families had low levels of education, the characteristics of their neighbours were strongly related to the chance that they would attend university. The third paper found that a post-secondary preparatory course that encouraged students to enrol in advanced classes and supported them academically could in the short-run help students achieve credit for an advanced math course. The students who were encouraged to take the math class were more likely to come from single-parent families, and had themselves lower educational aspirations.

The role of gender is one aspect of the research I present in this collection of papers that was beyond the scope of the project. Differences in outcomes and impacts across gender did emerge as interesting findings. For example, in the first paper, girls were less likely to drop-out of high school and less affected by parental aspirations. In the last paper, boys were more likely to be among the students who were encouraged to take an advanced math class by a program offering a post-secondary preparatory course. While there is some recent research exploring gender differences in educational attainment, for example Frenette and Zeman (2007), the issues related to gender raised by these papers are excellent topics for future research.

Several themes have emerged from the preceding chapters which lead to at least three general conclusions, which can be drawn from the findings in this collection of papers. First, survey measures of family income have only limited capacity to explain variations in educational outcomes among Canadian youth. In both of the first two papers, after controlling for parental education and reading skills, family income had only small impacts on the chances of high school drop-out and university participation.

It is important to emphasize, however, that from the point of view of a policy maker not much can be inferred from this finding. Generally, permanent income is of more interest from the policy perspective because families are expected to make decisions based on their long-run expectations rather than short-run fluctuations in income. The measures of family income used in these papers are self-reported current income, which are without doubt measured with error.

Current income may matter if individuals and families face borrowing constraints or if interest rates depend on income. In this dissertation, while there is little positive evidence of their impact, neither is there any solid evidence against the effects of borrowing constraints. Some researchers infer evidence of constrained behaviour from the effects of current income. In this dissertation, such inference is not warranted. For example, although I find little effect of family income, I do find that reading skills at age 15 have a large effect on university participation at age 21. If families know they will face a borrowing constraint when their child reaches university-age they may choose not to make early investments in their cognitive development. Heckman has emphasized this notion in the literature, often referring to under-investments in early years as a long-run constraint (e.g. Carneiro and

Heckman 2003). This collection of papers presents ample evidence of long-run constraints.

The second conclusion that can be drawn from my dissertation is an example of a long-run constraint. This conclusion underscores the importance of educational aspirations. In the first two papers, a variable which asks parents to report the highest level of education they hope their child will achieve plays an important role in shaping educational outcomes. This variable potentially captures a number of factors including information parents have about their children's ability and motivation, and the perceived value of education whether that stems from market returns or consumption value. The second chapter of my dissertation used this variable to help separate the effects of ability from the effects of parental valuations of education. In the third chapter, the variable was used as a control and was found to have a large impact on university participation. In the fourth chapter, I found children's own aspirations played an important role. Specifically, I found that students who were encouraged to take an advanced math class were less likely to hope they would achieve a Bachelors degree before the program began.

Finally, my dissertation draws attention to the importance of parental education in determining youth's educational outcomes. In particular, the relationship between educational outcomes and having parents with at least a bachelors degree is staggering. Children whose parents both have a BA are virtually guaranteed to complete high school. The overwhelming majority of children from BA-families also attend university. Yet, the first paper in this dissertation suggests that the direct effects of parental education are minimal. What does this imply for policy?

From one perspective, these findings imply that policies aimed at shifting the one observable characteristic which is most strongly related to children's education will not be effective in improving their outcomes. However, that perspective views the issue from *within* a generation. Across generations, the results suggest that improvements in education within disadvantaged groups, for example African-

Americans or First Nations people, might be permanent. All of the factors that are strongly correlated with parental education, including higher skills and parental valuations of education, combine to provide children born into families with Bachelors degrees with an apparent advantage in obtaining education. Moreover, there is evidence in this dissertation that potentially negative factors have less deleterious effects on children from BA-families. For example, children whose parents' have a Bachelors degree are not affected by living in lower quality neighbourhoods. While these findings point to the long-run potential of education to improve economic outcomes among historically disadvantaged populations, they also caution that in the short-run the mechanisms by which policy makers can improve education are profoundly complex.

5.1 Bibliography

P. Carneiro and J. Heckman. The evidence on credit constraints in post–secondary schooling. *Economic Journal*, 482(705-734), 2003.

M. Frenette and K. Zeman. Why are most university students women? Evidence based on academic performance, study habits and parental influences. Analytical Studies Branch Research Paper Series. Catalogue No. 11F0019MIE-No. 303, Statistics Canada, 2007.

Appendix A

Supplementary Tables

Table A.1: Chapter 2 Sample Means (Standard errors in parenthesis)

	Boys	Girls
Dropping out and Measurements		
Dropping out	0.056 (0.003)	0.036 (0.002)
PISA Q1	0.268 (0.005)	0.161 (0.004)
PISA Q2	0.252 (0.005)	0.223 (0.005)
PISA Q3	0.252 (0.005)	0.279 (0.005)
PISA Q4	0.228 (0.005)	0.337 (0.005)
Grades (Less than 59)	0.098 (0.003)	0.056 (0.003)
Grades (60 to 69)	0.211 (0.005)	0.154 (0.004)
Grades (70 to 79)	0.361 (0.005)	0.323 (0.005)
Grades (80 and above)	0.329 (0.005)	0.468 (0.005)
Parpref	0.614 (0.005)	0.701 (0.005)
Saved	0.595 (0.006)	0.589 (0.005)
Getby	0.231 (0.005)	0.400 (0.005)
Hmwrk	0.205 (0.005)	0.299 (0.005)
Sample	7,755	8,375

continued, next page

Table A.2: Chapter 2 Sample Means continued (Standard errors in parenthesis)

	Boys	Girls
Control Variables		
Parents' highest educational attainment –Reference both parents have a BA or higher		
Both parents have BA	0.102 (0.003)	0.101 (0.003)
One parent has BA	0.182 (0.004)	0.178 (0.004)
At least one parent has PSE below BA	0.415 (0.006)	0.398 (0.005)
Both parents have a high school diploma	0.097 (0.003)	0.090 (0.003)
One parent has a H.S. diploma	0.111 (0.004)	0.125 (0.004)
Both parents have less than H.S.	0.093 (0.003)	0.108 (0.003)
Other background characteristics		
Aboriginal	0.030 (0.002)	0.026 (0.002)
Immigrant	0.082 (0.003)	0.087 (0.003)
Rural	0.237 (0.005)	0.248 (0.005)
Number of moves	2.059 (0.027)	2.088 (0.026)
Number of siblings	1.346 (0.011)	1.335 (0.011)
Local youth unemployment rate	13.750 (0.069)	13.942 (0.068)
*Marginal effect of a difference in unemployment rate from top and bottom quartiles		
Family structure–Reference two biological parent families		
Other two parent families	0.123 (0.004)	0.121 (0.004)
Lone parent family	0.140 (0.004)	0.167 (0.004)
Sample	7,755	8,375

Source: Youth in Transition Survey, Cycle 3 (Cohort A)

Table A.3: Selected parameters from unobserved factor system estimators (Standard errors in parenthesis)

	1	2
Intercept	-1.433 (0.441)***	-1.026 (0.486)**
$\lambda_{d\theta_1}$	-0.011 (0.001)***	-0.009 (0.001)***
$\lambda_{d\theta_2}$	-0.355 (0.141)**	-0.410 (0.107)***
$\lambda_{d\upsilon_p}$	-4.348 (1.571)***	-2.025 (0.448)***
Reading Scores (*PISA*)		
Intercept	510.918 (5.322)***	536.624 (4.900)***
$\lambda_{T\theta_1}$	1	1
Variance parameter $\ln\left(\sigma^2_{u_1}\right)$	4.151 (0.026)***	3.944 (0.048)***
Parent's Aspirations (*Parpref*)		
Intercept	0.166 (0.101)	0.061 (0.086)
$\lambda_{p\theta_1}$	0.012 (0.001)***	0.009 (0.001)***
$\lambda_{p\theta_2}$	-0.139 (0.144)	0.015 (0.058)
$\lambda_{p\upsilon_p}$	2.482 (0.858)***	1.199 (0.188)***
Not wanting to just get by (*getby*)		
Intercept	-3.156 (0.281)***	-3.096 (0.210)***
$\lambda_{c\theta_2}$	1	1
$\lambda_{c\upsilon_p}$	11.898 (3.937)***	1.959 (0.289)***
Factor distribution vary with parents' education	N	Y
Sample	7,755	7,755

continued, next page

Table A.4: Selected parameters from unobserved factor system estimators continued (Standard errors in parenthesis)

	1	2
Grades (*grd*)		
Intercept	59.288 (1.190)***	66.352 (1.130)***
$\lambda_{g\theta_1}$	0.115 (0.005)***	0.090 (0.004)***
$\lambda_{g\theta_2}$	13.841 (1.949)***	7.520 (0.529)***
$\lambda_{g\upsilon_p}$	39.364 (12.121)***	10.349 (1.516)***
Variance parameter $\ln\left(\sigma^2_{u_4}\right)$	1.464 (0.066)***	1.872 (0.026)***
Save for children's education (*saved*)		
Intercept	-1.428 (0.185)***	-1.300 (0.194)***
$\lambda_{s\theta_1}$	0.002 (0.000)***	0.000 (0.000)
$\lambda_{s\theta_2}$	-0.079 (0.083)	-0.213 (0.059)***
$\lambda_{s\upsilon_p}$	1	1
Complete home work on time (*hmwrk*)		
Intercept	-3.156 (0.281)***	-3.096 (0.210)***
$\lambda_{h\theta_2}$	1.767 (0.257)***	1.764 (0.163)***
$\lambda_{h\upsilon_p}$	5.644 (1.764)***	0.849 (0.148)***
Factor locations		
θ^H_{11}	95.493 (2.884)***	95.019 (3.428)***
θ^L_{11}	-84.558 (3.773)***	-96.942 (2.868)***
θ^H_{12}	0.836 (0.110)***	1.267 (0.068)***
υ^H_p	0.196 (0.059)***	0.754 (0.081)***
Factor distribution vary with parents' education	N	Y
Sample	7,755	7,755

Appendix B

Likelihood Function

In this appendix, we present an example contribution to the factor likelihood function for person *i* who is a dropout; has a test score in the lowest quartile; has parents who state they hope their child gets a BA; states he just wants to get by in effort; has an average overall grade below 59; has parents who saved for their education; and hands in homework late.

$$\begin{aligned}
\textstyle\sum_j \sum_k \sum_m p^j_{\theta_{11}} p^k_{\theta_{12}} p^m_{\upsilon_p} \quad & F(-z\gamma - \lambda_{d\theta 1}\theta^j_{11} - \lambda_{d\theta 2}\theta^k_{12} - \lambda_{d\upsilon}\upsilon^m_p) \\
& F(PISA_1 - x_1\delta_1 - \theta^j_{11}) \\
& F(-x_2\delta_2 - \lambda_{p\theta 1}\theta^j_{11} - \lambda_{p\theta 2}\theta^k_{12} - \lambda_{p\upsilon}\upsilon^m_p) \\
& F(-x_3\delta_3 - \theta^k_{12} - \lambda_{c\upsilon}\upsilon^m_p) \\
& F(59 - x_4\delta_4 - \lambda_{g\theta 1}\theta^j_{11} - \lambda_{g\theta 2}\theta^k_{12} - \lambda_{g\upsilon}\upsilon^m_p) \\
& F(-x_5\delta_5 - \lambda_{s\theta 1}\theta^j_{11} - \lambda_{s\theta 2}\theta^k_{12} - \upsilon^m_p) \\
& F(-x_6\delta_6 - \lambda_{h\theta 2}\theta^k_{12} - \lambda_{c\upsilon}\upsilon^m_p)
\end{aligned} \tag{B.1}$$

where: the F(·)'s are cumulative normal distribution functions; j, k and m index the points of support in the θ_{11}, θ_{12} and υ_p distributions, respectively; the p's are probabilities associated with the points of support; z corresponds to the vector of all observable covariates in the dropout equation with a vector of associated coefficients, γ; and $x_1 \ldots x_6$ are the vectors of observable covariates in the measurement equations with vectors of associated coefficients, $\delta_1 \ldots \delta_6$. In our data, dropping out is a binary variable as are $parpref$ (a dummy equalling one if the parents hope their child will obtain a BA or more and zero otherwise), *saved* (which takes a value of one if the parents saved for their child's education), *getby* (which equals 1 if children say they just want to get by in terms of effort), and *hmwork* (which equals 1 if the child always completes his assignments). These variables contribute simple Probit type expressions to the likelihood conditional on the factor values. We divide the PISA test scores into quartiles and use indicators for the quartile of the *PISA* variable. Thus, the contribution to the likelihood function is in the form of components of an ordered Probit. Here, $PISA_1$ is the test score value that defines the upper bound of the first quartile. Similarly, we group grades in four categories (59 and less, 60 to 69, 70 to 79, and 80 and above). As a result, the contributions for this variable also take the form of ordered Probit expressions.