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

IFS Working Paper W09/21

Kelly Foley
Giovanni Gallipoli
David A. Green

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

Kelly Foley, UBC
Giovanni Gallipoli, UBC and RCEA
David A. Green, UBC and IFS *

September 9, 2009

Abstract

We use a large, rich Canadian micro-level dataset 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 factor model set out in Carneiro, Hansen, and Heckman (2003). Specifically, we consider the impact of cognitive and non-cognitive ability and the value that parents place on education. Our results support three main conclusions. First, cognitive 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 teenager 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. These effects are estimated while conditioning on ability at age 15. Thus, under some assumptions, they reflect parental influences during the upper teenage years and are in addition to any impact they might have in the early childhood years leading up to age 15. Third, parental education has no direct effect on dropping out once we control for ability and parental valuation of education. Overall, 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. Our work also provides a small methodological contribution by extending the standard factor based estimator to allow a more non-linear relationship between the factors and a co-variate of interest. We show that allowing for non-linearities has a substantial impact on estimated effects.

---

*Correspondence: Department of Economics, University of British Columbia, 997-1873 East Mall, Vancouver BC, V6T 1Z1 Canada. The statistical analysis in this paper was produced from Statistics Canada microdata. The interpretation and opinions expressed are those of the authors and do not represent those of Statistics Canada. We thank participants at the CLSRN and IRP workshops at the University of Toronto and University of Wisconsin, respectively, for their comments.

1 Introduction

The strong correlation between family socio-economic status [1] and dropping out of high school is well known (e.g., Eckstein and Wolpin, 1999 for the U.S. and Belley, Frenette, and Lochner, 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 result in a social efficiency loss. Second, understanding this correlation is important when thinking about redistribution.[2] Our goal in this paper 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, Haller, and Portes, 1969, Alexander, Eckland, and Griffin, 1975 and Haveman, Wolfe, and Spaulding, 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 would call for policy responses which can supply for parental and family influences. If, instead, the answers reflect insider knowledge about a child's

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

[2]Under one theory of justice (Dworkin 1981a, 1981b) 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. Understanding 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).

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 outcomes.[3]

Cunha and Heckman, (2007, 2008) and Cunha, Heckman, and Schennach (2006) – hereafter, CH07, CH08, and CHS, respectively – investigate the production of both cognitive and non-cognitive abilities using dynamic factor models. As in Heckman, Stixrud, and Urzua (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 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 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, Hansen, and Heckman (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

[3]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.

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 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 a shape for factor distributions which is 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, Meghir, and Parey (2007).[4]

The paper is organized as follows. In Section 2 we present a simple life-cycle model describing the decision to drop out from high school. In Section 3 we map the model to data, setting out an estimable counterpart. In Section 4 we describe the data, and Section 5 contains results from the various estimators described earlier. In Section 6 we summarize and conclude.

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.[5]

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 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

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

[5] 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.

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.

2.1 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,

$$\theta_0 = f_1(\theta_F, \iota) \tag{1}$$

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,

   $$PE = f_2(\theta_F, \nu_p, \eta) \tag{2}$$

   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,

   $$\theta_1 = f_3(\theta_0, PE, \nu_p) \tag{3}$$

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 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}). [6] 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,

$$w_2^p = \alpha_0^p + \alpha_1\theta_1 + \alpha_2 grd \tag{4}$$

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 youths[7], by specifying,

$$\alpha_0^p = \alpha_{01} + \alpha_{02}PE + \alpha_{03}\upsilon_p \tag{5}$$

where υ_p corresponds to parental predictions about returns to education.[8] 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.

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

[7] Notice that w_2^p assumes values on some interval $[\underline{\mathrm{w}},\bar{\mathrm{w}}]$ and that predicted education outcomes are approximated, at age 15, by equation (4) .

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

The academic performance measure in equation (4) is determined by,

$$grd = \psi_0 + \psi_1 e + \psi_2 PE + \psi_3 x_s + \psi_4 \theta_1 + \psi_5 \upsilon_p \tag{6}$$

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 (4) and (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 (4) can be seen as a linearization of such processes.[9]

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 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).[10] 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. [11]

[9]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.

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

[11]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.

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

$$V_S = \max_e g(e) + f_S(PI, FI) + \beta n w_2^p \tag{7}$$

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

$$V_W = \max_e g(e) + f_W(PI, FI) + (1 + \beta n) w_{LHS} \tag{8}$$

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

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

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

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

$$e^*_{1-d} = \beta n \psi_1 \alpha_2 - \gamma \tag{9}$$

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[12]. It will also be affected by parental valuations of the returns to education and peer characteristics through their impacts on γ.

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

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

income, PI, can itself be approximated as a linear function of parental education, this difference is given by,

$$I_D = \gamma_0 + \gamma_1 FI + \gamma_2 PE + \gamma_3 w_{LHS} + \gamma_4 x_s + \gamma_5 z + \lambda_{d\theta 1}\theta_{11} + \lambda_{d\theta 2}\theta_{12} + \lambda_{dv} v_p + u_0 \quad (10)$$

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 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 (10) is complicated by the fact that we do not directly observe θ_{11}, θ_{12} or v_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.

3 Empirical Strategies

3.1 Reduced Form

The simplest approach to estimating (10) is to ignore θ_{11}, θ_{12} and v_p and implement (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.

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,

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

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 (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.[13]

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,

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

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, Duckworth, Heckman, and ter Weel (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 factor, 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

[13]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.

to get by," is true for him or her.[14] We code a variable equaling 1 if they answer 'Never' and assume this is determined by an underlying index function,

$$getby = \delta_{30} + \delta_{31}PE + \lambda_{c\theta 2}\theta_{12} + \lambda_{cv}\upsilon_p + u_3 \tag{13}$$

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.[15]

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 (11) for θ_{11} and substitute into (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 (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.

3.3 Unobserved Factor System Estimators

3.3.1 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

[14] 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 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.

[15] 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.

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, (10). The test score equation (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 (6), yielding,

$$grd = \delta_{40} + \delta_{41}PE + \delta_{42}x_s + \delta_{43}z + \lambda_{g\theta 1}\theta_{11} + \lambda_{g\theta 2}\theta_{12} + \lambda_{g\upsilon}\upsilon_p + u_4 \tag{14}$$

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,

$$saved = \delta_{50} + \delta_{51}PE + \delta_{52}FI + \lambda_{s\theta 1}\theta_{11} + \lambda_{s\theta 2}\theta_{12} + \lambda_{s\upsilon}\upsilon_p + u_5 \tag{15}$$

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,

$$hmwork = \delta_{60} + \delta_{61}PE + \lambda_{h\theta 2}\theta_{12} + \lambda_{h\upsilon}\upsilon_p + u_6 \tag{16}$$

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

Together, equations (10) through (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), λ_{sv} (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.[16] 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.[17]

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 dropout process) is given by,

$$\lambda_{d\theta 1} = \frac{cov(I*_G, T*)}{cov(grd*, T*)} \tag{17}$$

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 (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.

[16] 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.

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

The result in (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.[18] 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, assuming that equations (1) through (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.

3.3.2 Extended Estimator

The assumption that all relevant equations are linear in the underlying factors is potentially strong. In particular, the results in Gallipoli, Meghir, and Violante (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, (3). To understand the implication of these non-linearities for our estimation, note that we are interested in characterizing the conditional distribution,

$$p(Y, \theta | PE; \Gamma) = \int p(Y | \theta, PE; \Gamma) p(\theta | PE; \Gamma) d\theta \tag{18}$$

where Y is the matrix of outcomes (both dropping out and the measurement values), θ is a

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

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 (1)–(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 (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.[19]

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: in principle, two students with the same abilities and parental valuation might have similar probability of dropping out even if their parents have very different education levels. Notice, however, that 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.[20]

[19] To verify this guess, we constructed a Monte Carlo exercise in which we generated values for a single factor and for PE based on equations (1)–(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 proxies for the factor, measured with error. When we specify (1)–(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 (1)–(3) to be nonlinear, the proxy, IV, and standard system estimator all produced upwardly biased estimates of the PE coefficient.

[20] 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,

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 cognitive 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, Duckworth, Heckman, and ter Weel, 2008). Based on this (and given how important parental valuation turns out to be in the dropout process), it seems unlikely to us that ν_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 ν_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 ν_p factor as capturing parental valuation of education.

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.

we present our results using that specification and omit the correlated factor estimates to save on space.

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.[21] In the second stage, students were sampled within the 1,187 schools.[22] The sample size within each school was chosen to facilitate school-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.[23]

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

[21] 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.

[22] Schools were excluded from the sample if fewer than 3 students were present or likely to respond 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.

[23] 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.

not already dropped out or chosen to graduate after 12 years, are in the last terms of high school, and are likely interested 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 [24] and obtained very similar results to those presented here. Thus, we do not believe these two anomalies are driving any of our results. [25]

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, Frenette, and Lochner (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.[26] Finally, to place Canada's experience in context, Belley, Frenette,

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

[25]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.

[26]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 effectively 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

and Lochner (2008) 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.

5 Results

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 graduates; 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 C1 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,

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.

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 on 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.[27] 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 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, 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 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, Frenette, and Lochner (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.

[27] 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

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, Frenette, and Lochner (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 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 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 1, we plot the probabilities associated with different parental education levels from this specification.[28] 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.

[28]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 1 correspond to families where both parents have the specified education level.

Table 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 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	0.028	0.035	0.034
		(0.030)*	(0.021)	(0.025)	(0.023)
Immigrant		-0.027	-0.019	-0.014	-0.028
		(0.014)*	(0.011)*	(0.014)	(0.011)**
Rural		-0.010	-0.016	-0.016	-0.013
		(0.009)	(0.007)**	(0.008)**	(0.007)*
Number of moves		0.003	0.003	0.003	0.003
		(0.002)**	(0.001)**	(0.001)**	(0.001)**
Age in months		0.003	0.003	0.003	0.003
		(0.002)**	(0.001)**	(0.001)**	(0.001)**
Number of siblings		-0.006	-0.003	-0.004	-0.004
		(0.004)	(0.003)	(0.004)	(0.004)
Family structure –Reference two biological parent families					
Other two parent families		0.080	0.052	0.063	0.061
		(0.022)***	(0.016)***	(0.019)***	(0.018)***
Lone parent family		0.008	0.012	0.011	0.011
		(0.012)	(0.010)	(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
Family income is the before-tax family income divided by the square root of the number of household members.
Other two parent families include biological and adoptive parents, step-parents and guardians.
The reference person is a non-Aboriginal non-immigrant youth who lives in an urban area in Ontario, living with two biological parents who both have a Bachelors degree and do not hope their child obtains a university degree. The reference youth scored in the top quartile on the PISA reading test.
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: 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: 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 (0.020)	0.022 (0.015)	0.022 (0.014)*	0.021 (0.016)
Immigrant		-0.020 (0.007)***	-0.015 (0.005)***	-0.017 (0.006)***	-0.017 (0.006)***
Rural		-0.001 (0.005)	-0.002 (0.004)	-0.003 (0.005)	-0.002 (0.004)
Number of moves		0.002 (0.001)**	0.001 (0.001)**	0.002 (0.001)**	0.002 (0.001)**
Age in months		0.002 (0.001)**	0.001 (0.001)**	0.002 (0.001)**	0.002 (0.001)**
Number of siblings		-0.004 (0.002)*	-0.003 (0.002)*	-0.003 (0.002)	-0.003 (0.002)*
Family structure –Reference two biological parent families					
Other two parent families		0.028 (0.012)**	0.016 (0.008)*	0.022 (0.010)**	0.019 (0.009)**
Lone parent family		0.009 (0.008)	0.009 (0.007)	0.009 (0.008)	0.010 (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
Family income is the before-tax family income divided by the square root of the number of household members.
Other two parent families include biological and adoptive parents, step-parents and guardians.
The reference person is a non-Aboriginal non-immigrant youth who lives in an urban area in Ontario, living with two biological parents who both have a Bachelors degree and do not hope their child obtains a university degree. The reference youth scored in the top quartile on the PISA reading test.
Marginal effects are evaluated for a youth whose parents both have a high school diploma at the mean of all the other variables.

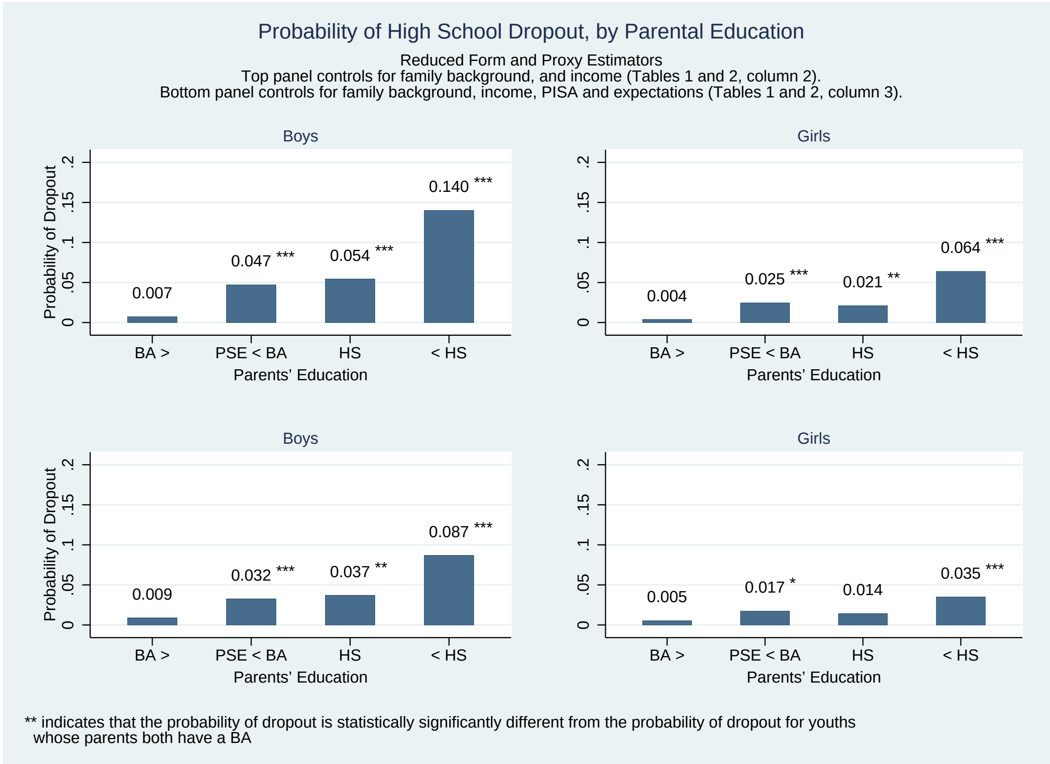

Figure 1: Probability of dropping out by parental education.

The results for females in column 2 of Table 2 and plotted in the top right graph in Figure 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.

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 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'.[29] 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.

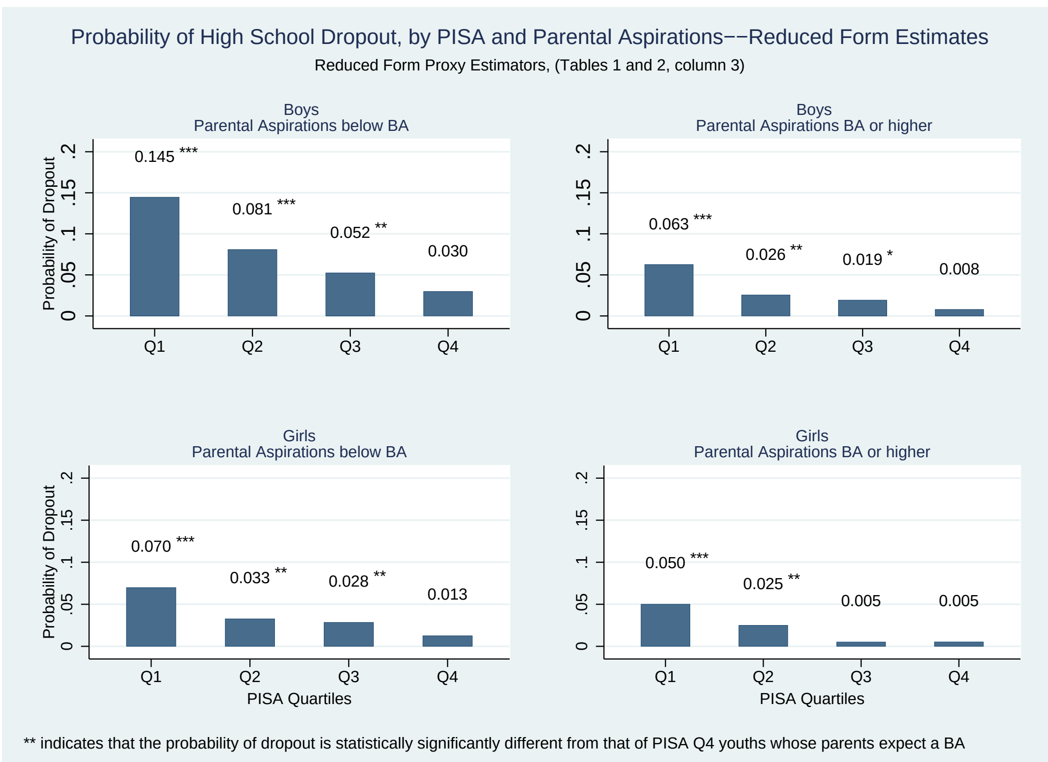

Figure 2: Probability of dropping out by PISA and parental aspirations – Reduced Form.

We introduce the PISA quartiles interacted with parental aspirations in column 3 of Tables 1 and 2. We also plot the predicted probabilities of dropping out (evaluated at the mean) for the

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

various categories in Figure 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 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 probability 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 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 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 1 and 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.

Table 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	-0.002	-0.002	-0.002	-0.001
	(0.001)***	(0.001)**	(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	0.015	0.011	0.012	0.009
	(0.008)*	(0.008)*	(0.006)*	(0.006)*	(0.006)
At least one parent has PSE below BA	0.024	0.023	0.020	0.020	0.016
	(0.007)***	(0.007)***	(0.006)***	(0.006)***	(0.006)***
Both parents have a high school diploma	0.028	0.027	0.023	0.024	0.017
	(0.011)***	(0.010)***	(0.009)**	(0.009)***	(0.008)**
One parent has a H.S. diploma	0.032	0.030	0.027	0.024	0.023
	(0.011)***	(0.010)***	(0.009)***	(0.009)***	(0.010)**
Both parents have less than H.S.	0.078	0.075	0.071	0.064	0.065
	(0.017)***	(0.016)***	(0.015)***	(0.015)***	(0.016)***
PISA scores and parents' aspirations– Reference PISA Quartile 4 and BA aspirations					
Below BA aspirations–PISA Quartile 1	0.137	0.132	0.102	0.091	0.073
	(0.026)***	(0.026)***	(0.022)***	(0.020)***	(0.019)***
Below BA aspirations–PISA Quartile 2	0.073	0.069	0.051	0.050	0.035
	(0.020)***	(0.019)***	(0.015)***	(0.016)***	(0.014)***
Below BA aspirations–PISA Quartile 3	0.045	0.042	0.031	0.033	0.031
	(0.018)**	(0.017)**	(0.014)**	(0.014)**	(0.014)**
Below BA aspirations–PISA Quartile 4	0.022	0.020	0.019	0.017	0.006
	(0.015)	(0.014)	(0.014)	(0.014)	(0.008)
BA and above aspirations–PISA Quartile 1	0.055	0.051	0.034	0.039	0.037
	(0.016)***	(0.015)***	(0.011)***	(0.013)***	(0.013)***
BA and above aspirations–PISA Quartile 2	0.018	0.017	0.012	0.012	0.010
	(0.007)**	(0.007)**	(0.006)*	(0.006)*	(0.006)*
BA and above aspirations–PISA Quartile 3	0.011	0.011	0.009	0.009	0.010
	(0.007)*	(0.007)*	(0.006)	(0.006)	(0.006)*
Non-cognitive outcomes					
Child never 'just wants to get by'		-0.024	-0.012	-0.011	-0.010
		(0.008)***	(0.007)*	(0.007)	(0.007)
Always does homework on time			-0.024	-0.023	-0.020
			(0.007)***	(0.007)***	(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 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
Family income is the before-tax family income divided by the square root of the number of household members.
Low educational resources takes on the value one if a high school principal reported that the learning of grade 10 students is hindered by the lack of instructional material a lot or to some extent.
The reference person is a non-Aboriginal non-immigrant youth who lives in an urban area in Ontario, living with two biological parents who both have a Bachelors degree and do not hope their child obtains a university degree. The reference youth scored in the top quartile on the PISA reading test.
Marginal effects are evaluated for a youth whose parents both have a high school diploma at the mean of all the other variables.

Table 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 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
Family income is the before-tax family income divided by the square root of the number of household members.
Low educational resources takes on the value one if a high school principal reported that the learning of grade 10 students is hindered by the lack of instructional material a lot or to some extent.
The reference person is a non-Aboriginal non-immigrant youth who lives in an urban area in Ontario, living with two biological parents who both have a Bachelors degree and do not hope their child obtains a university degree. The reference youth scored in the top quartile on the PISA reading test.
Marginal effects are evaluated for a youth whose parents both have a high school diploma at the mean of all the other variables.

5.2.1 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 1 and 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 3 and 4 reproduce column 3 of Tables 1 and 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 are very similar after including the proxy for non-cognitive skills.

In the next column of Tables 3 and 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. [30]

In the final column of Tables 3, we incorporate school characteristics that 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.[31]

The last column of Table 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

[30] 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.

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

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.

5.3 Factor Estimators

In this section, we present results from the full factor models set out in Section 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.[32] Thus, we implement a specification with three points of support for cognitive ability and two each for parental valuation and non-cognitive ability. Table 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 1 for the proxy estimator.

[32]Specifically, the model returned a probability mass for the additional points of support very close to zero and imprecisely estimated.

Table 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)
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
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.
The reference person is a non-Aboriginal non-immigrant youth who lives in an urban area in Ontario, living with two biological parents who both have a Bachelors degree.
Marginal effects are evaluated for a youth whose parents both have a high school diploma at the mean of all the other variables.

Once again, the impact of family income is essentially zero: notably smaller than the already small impacts in Belley, Frenette, and Lochner (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 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 (11) is relatively small.

Estimates of the factor loads, locations and associated probabilities are given in Table $A2$. The factor loads indicate the 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 6A, 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 6A and 6B are also shown in Figures 3 and 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 6A shows the fitted dropout probabilities evaluated at high non-cognitive ability and the bottom panel shows the same for low non-cognitive ability.

Table 6A: 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)
Standard errors clustered by high school.
Evaluated at the mean of family income and other socio-economic background variables.

Table 6B: 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)
Standard errors clustered by high school.
Evaluated at the mean of family income and other socio-economic background variables.

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 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.[33]

The estimated marginal impacts from the extended factor model are given in the second column of Table 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 6b, 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.[34] 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 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.

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

[34]Note that the small and insignificant differences in marginal impacts of parental education in Table 5 are converted into larger (though still statistically insignificant) differences in Table 6b because the curvature of the normal cumulative distribution function toward the tails converts small parameter differences into larger probability differences.

Table 7A: 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)
Standard errors clustered by school.

Table 7B: 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)

Table 7A shows the estimated distribution of teenagers across the points of support for each of the factors and each parental education level. Table 7B 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 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 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. The 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, Meghir, and Parey (2007) find significant effects of maternal

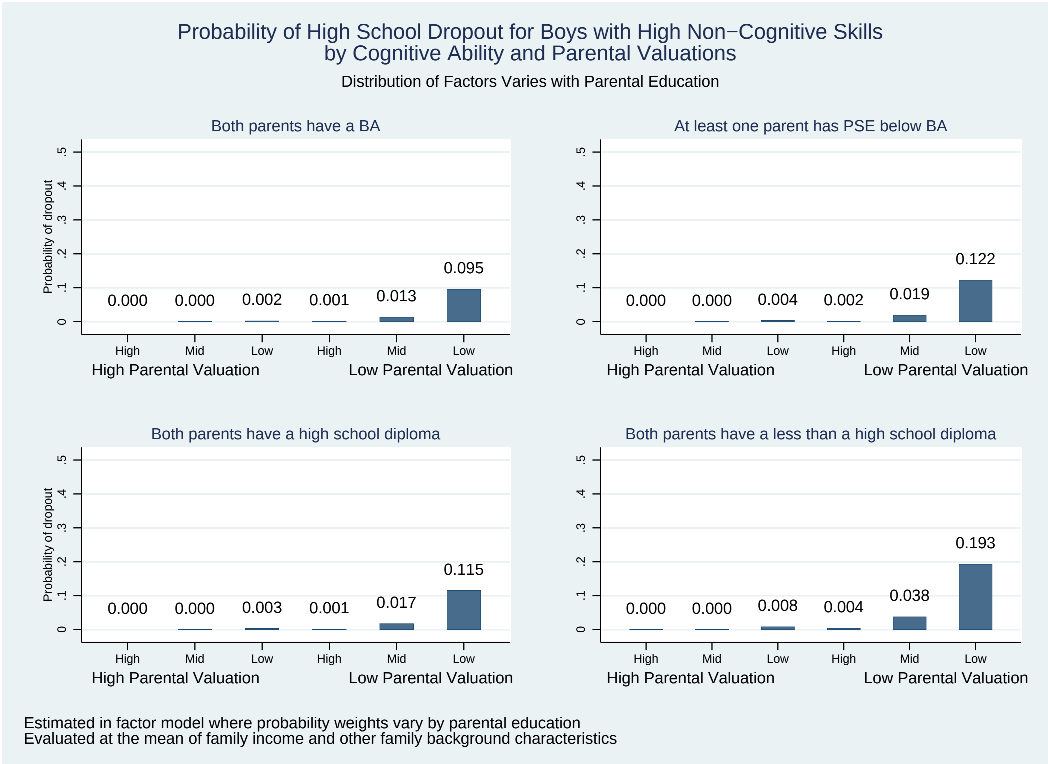

Figure 3: Probability of dropping out by cognitive ability and parental valuations for boys with high non-cognitive skills – extended factor model.

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, Meghir, and Parey, 2007 find is the case), the estimated parental education effects identified by these papers may ultimately occur through the channel we identify.

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 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

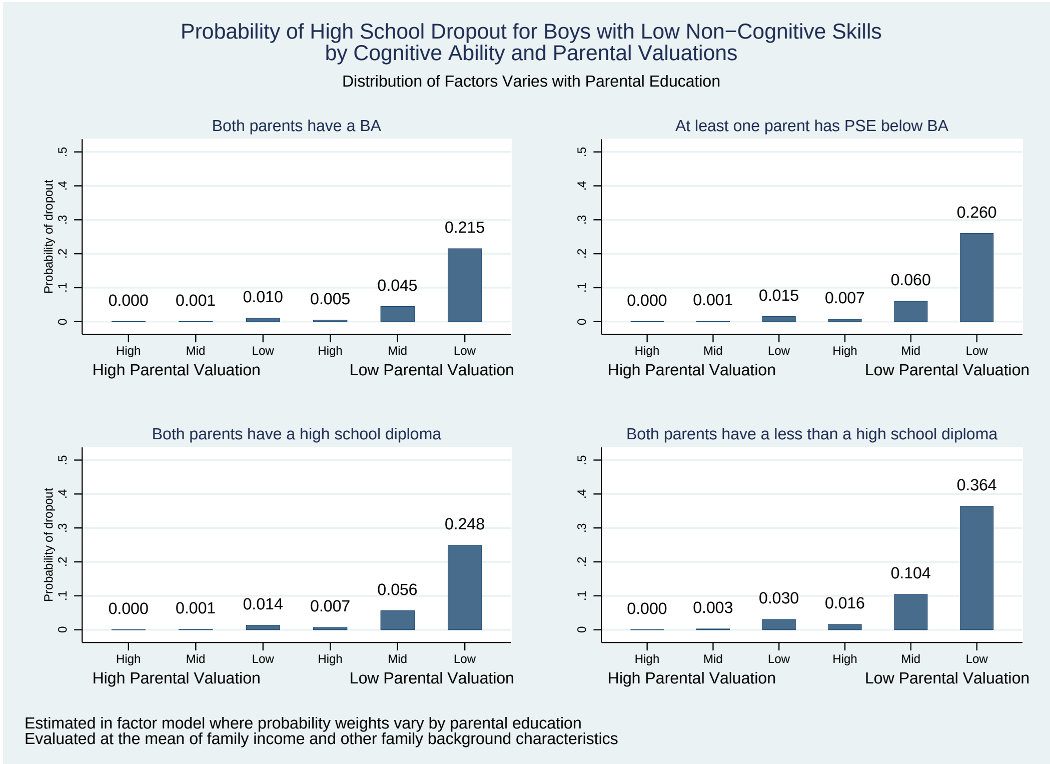

Figure 4: Probability of dropping out by cognitive ability and parental valuations for boys with low non-cognitive skills – extended factor model.

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, Haller, and Portes, 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 be reflected in cognitive skills and family valuations of education.

We investigate this finding in more detail using an unobserved factor approach based on Carneiro, Hansen, and Heckman (2003). Given arguments in Cunha, Heckman, and Schennach (2006) that ability production functions may be non-linear in parental investments, we consider an extension to the estimator in Carneiro, Hansen, and Heckman (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 span over the very long-term, 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, it seems that 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 provides some justification for policy intervention.

A Appendix Tables

Table A1: 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 A1: Sample means (cont'd)
(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 A2: 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 A2: Selected parameters From unobserved factor system estimators (cont'd)
(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

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.

$$\begin{aligned}\sum_j \sum_k \sum_m p^j_{\theta_{11}} p^k_{\theta_{12}} p^m_{\upsilon_p} \quad & F(-z\gamma - \lambda_{d\theta 1}\theta^j_{11} - \lambda_{d\theta 2}\theta^k_{12} - \lambda_{d\upsilon}\upsilon^m_p) \\ & F(PISA_1 - x_1\delta_1 - \theta^j_{11}) \\ & F(-x_2\delta_2 - \lambda_{p\theta 1}\theta^j_{11} - \lambda_{p\theta 2}\theta^k_{12} - \lambda_{p\upsilon}\upsilon^m_p) \\ & F(-x_3\delta_3 - \theta^k_{12} - \lambda_{c\upsilon}\upsilon^m_p) \\ & F(59 - x_4\delta_4 - \lambda_{g\theta 1}\theta^j_{11} - \lambda_{g\theta 2}\theta^k_{12} - \lambda_{g\upsilon}\upsilon^m_p) \\ & F(-x_5\delta_5 - \lambda_{s\theta 1}\theta^j_{11} - \lambda_{s\theta 2}\theta^k_{12} - \upsilon^m_p) \\ & F(-x_6\delta_6 - \lambda_{h\theta 2}\theta^k_{12} - \lambda_{c\upsilon}\upsilon^m_p)\end{aligned} \tag{19}$$

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 $x_1 ... x_6$ are the vectors of observable covariates in the measurement equations with vectors of associated coefficients, $\delta_1 ... \delta_6$. 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.

References

Alexander, K., B. Eckland, and L. Griffin (1975): "The Wisconsin Model of Socioeconomic Achievement: A Replication," *The American Journal of Sociology*, 81(2), 324–342.

Attanasio, O., and K. Kaufmann (2009): "Educational Choices, Subjective Expectations, and Credit Constraints," NBER Working Papers 15087, National Bureau of Economic Research, Inc.

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

Belley, P., M. Frenette, and L. Lochner (2008): "Educational Attainment by Parental Income: A Canada - U.S. Comparison," mimeo, Dept of Economics, University of Western Ontario.

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. ter Weel (2008): "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, 43(4).

Carneiro, P., K. T. Hansen, and J. J. Heckman (2003): "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.

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

Carneiro, P., C. Meghir, and M. Parey (2007): "Maternal Education, Home Environments and the Development of Children and Adolescents," Wp15/07, Institute for Fiscal Studies.

Chamberlain, G. (1977): "Education, Income and Ability Revisited," *Journal of Econometrics*, 5(2), 241–257.

Coelli, M. (2009): "Parental Job Loss, Income Shocks and the Education Enrolment of Youth," Department of Economics Working Paper 1060, The University of Melbourne.

Cunha, F., and J. Heckman (2007): "The Technology of Skill Formation," *American Economic Review*, 97(2), 31–47.

Cunha, F., J. Heckman, and S. Schennach (2006): "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," mimeo, Dept of Economics, University of Chicago.

Cunha, F., and J. J. Heckman (2008): "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources*, 43(4).

Davies, M., and D. B. Kandel (1981): "Parental and Peer Influences on Adolescents' Educational Plans: Some Further Evidence," *Review of Economic Studies*, 87(2), 363–387.

de Broucker P. (2005): "Without a Paddle: What to do About Canada's Young Drop-outs," Report w/30, Canadian Policy Research Network.

Dworkin, R. (1981a): "What is Equality? Part 1: Equality of Welfare," *Philosophy and Public Affairs*, 10, 185–246.

——— (1981b): "What Is Equality? Part 2: Equality of Resources," *Philosophy and Public Affairs*, 10, 283–345.

Eckstein, Z., and K. I. Wolpin (1999): "Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities," *Econometrica*, 67(6), 1295–1340.

Gallipoli, G., C. Meghir, and G. Violante (2009): "Equilibrium Effects of Education Policies: a Quantitative Evaluation," UCL Working paper.

Haveman, R., B. Wolfe, and J. Spaulding (1991): "Childhood Events and Circumstances Influencing High School Completion," *Demography*, 28(1), 133–157.

Heckman, J., and B. Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52(2), 271–320.

Heckman, J. J., and L. Lochner (2000): "Rethinking Education and Training Policy: Understanding the Sources of Skill Formation in a Modern Economy," in *Securing the Future: Investing in Children from Birth to College*, ed. by S. Danziger, and J. Waldfogel, pp. 47–83. New York: Russell Sage Foundation.

Heckman, J. J., J. Stixrud, and S. Urzua (2006): "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior," *Journal of Labor Economics*, 24(3), 411–482.

Junor, S., and A. Usher (2003): "The Price of Knowledge 2004: Access and Student Finance in Canada," Canadian Millennium Scholarship Foundation.

Pintrich, P., and E. D. Groot (1990): "Motivational and Self-Regulated Learning Components of Classroom Academic Performance," *Journal of Educational Psychology*, 82(1), 33–40.

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

Sewell, W., O. Haller, and A. Portes (1969): "The Educational and Early Occupational Attainment Process," *American Sociology Review*, 34, 82–91.

Todd, P. E., and K. I. Wolpin (2006): "The Production of Cognitive Achievement in Children: Home, School and Racial Test Score Gaps," University of Pennsylvania, Working Paper.