# The gender gap in university enrollment: What role do skills and parents play?

Kelly Foley*
University of Saskatchewan

Draft: September 28, 2017

**Abstract**

Young men are far less likely than women to attend university across most OECD countries. Recent research also suggests that boys are also falling behind in their grades and educational aspirations during high school. Both grades and aspirations reflect many different individual characteristics and socio-economic circumstances. To uncover the deeper determinants of the gender gap in university participation, I use Canadian data from the Youth in Transition Survey to estimate a factor model based on a framework developed by Foley, Gallipoli, and Green (2014). I use that model to identify and quantify the impact of three factors: cognitive skills, non-cognitive skills and parental valuations of education (PVE). I find that all three factors play an important role in explaining both the level and the gap in university participation. The factor structure as a whole accounts for 64 percent of the gender gap, and the distribution of the PVE factor accounts for 23 percent. This result suggests that parents play a larger role than what is implied by decompositions employing only observed determinants.

*The analysis presented in this paper was conducted at the Saskatchewan Research Data Centre (SKY-RDC) which is part of the Canadian Research Data Centre Network (CRDCN). The services and activities provided by the SKY-RDC are made possible by the financial or in-kind support of the SSHRC, the CIHR, the CFI, Statistics Canada, and the University of Saskatchewan. The views expressed in this paper do not necessarily represent the CRDCNs or that of its partners.

# 1 Introduction

In most OECD countries, women enroll in university at higher rates than men. Using data from the United States, economists have proposed several explanations for such gender gaps in post-secondary schooling outcomes and academic achievement more generally. These include gender differences in non-cognitive skills (Jacob, 2002; Becker et al., 2010; Conger and Long, 2010), job opportunities and the returns to schooling, (Goldin, 1995), as well as, aspirations and plans for the future (Fortin et al., 2015).

Although gender gaps in university outcomes are often larger in Canada than in the U.S. (OECD, 2016), arguably less has been written about this topic using Canadian data. Frenette and Zeman (2007), who investigate the gender gap in university attendance using the Youth in Transition Survey (YITS), are a notable exception. The YITS is a longitudinal survey that follows a cohort of Canadian youth beginning when they were aged 15, and which also includes a parental survey. Frenette and Zeman (2007) perform a Oaxaca -Blinder decomposition to estimate how much of the gap can be attributed to various observed characteristics. Grades, scores on a reading test, study habits, and parental aspirations for their children's education are among the most important explanatory variables.

Grades and parental aspirations are themselves variables that reflect many different contributing factors. For example, when parents answer questions about how far they hope their children will go in education, their answers reflect not only their own hopes and values but also their knowledge of the child's skills and motivations. Additionally, the grades students achieve in high school depend on their own efforts and abilities, as well as help and encouragement from their parents. Interpreting the impact of these variables is further complicated because many of the important underlying factors are not directly observed.

In this paper, I take an additional step toward identifying and quantifying the underlying factors that explain the gender gap in university attendance among Canadian youth. In particular, my goal is to isolate any effects parents may have on university enrollment in the teenage years, beyond the stock of skills children have acquired up to that point. Like Frenette and Zeman (2007), I use the YITS.[1] In that data, 52% of girls and 37% of boys attended university by age 21, leading to a gap of 15 percentage points.

My approach to identification employs a model developed by Foley, Gallipoli, and Green (2014), hereafter FGG, which is based on the empirical methods introduced by Carneiro et al. (2003) and Cunha et al. (2005). In FGG, the model was used to estimate how much of the observed socioeconomic gradient in dropping out of high school among boys is explained by three unobserved factors. Following the literature emphasizing the importance and multi-dimensionality of skills (Cunha and Heckman, 2007; Cunha et al., 2010), FGG include cognitive and non-cognitive skills as factors. One of the contributions of FGG is to introduce a third orthogonal factor, interpreted as how much parents value education. I adapt the FGG model to the outcome of university enrollment, and examine the extent to which gender differences in the three unobserved factors explain the gap in enrollment.[2]

Variation in the parental valuation of education (PVE) factor is defined *conditional on cognitive and non-cognitive skills at age 15*, and is related to differences in parents' perceptions of the returns to university, of either a pecuniary or non-pecuniary nature, and differences in their willingness and ability to pay for their children's education. Empirically, the PVE factor is correlated with parents' reported aspirations for their children's education and whether they have saved for their child's education, holding family income constant.

While the factors are, by construction, orthogonal to each other, the interpretation of the third factor depends on the assumption that the two skills factors are a sufficient statistic for the stock of ability at age-15. In other words, the decision to enroll in university depends on the value of the skills factors and not on how those skills were acquired. Additionally, this requires that there are not other dimensions of skills, which both affect university enrollment and are

[1] I use one more cycle of the survey measuring participation at age 21 rather than 19. Because the data includes youth from Ontario, for whom a fifth year of high school was possible, participation rates increase between age 19 and 21.

[2] Because too few girls drop out of high school in the YITS data set, FGG were unable to explore gender differences.

orthogonal to the two included skills-factors. Under this assumption, the PVE factor reflects an influence parents exert on their teenagers from age 15 until the enrollment decision, which I measure at age 21.

I estimate the model allowing the distribution of the factors and their impact on outcomes to vary across gender. For both boys and girls, I find that all three factors have a significant impact on university enrollment. Cognitive skills have the largest impact, increasing the probability of attendance by more than .50 when comparing the lowest to the highest skill level. The impact of parental valuations varies across the two skill levels. Conditional on a medium level of both cognitive and non-cognitive skills, a high level of the PVE factor increases the probability of attendance by nearly 38 percentage points, for both boys and girls.

I also find that once I have controlled for the factors, allowing the distributions differ for males and females, the direct effect of gender, represented by an intercept shift, is no longer statistically significant. The predicted probability of attending university among those with the highest cognitive skill level is nearly identical for both boys and girls. Gender gaps do persist at some of the lower skill levels, however, the model predicts that a relatively small fraction of the data fall into those categories.

Overall, the results suggest that girls attend university more often than boys because they have higher levels of all three factors. The impact of cognitive skills on enrollment is also larger for girls, but there is no detectable gender difference in the impact of the other two factors. To quantify the importance of each factor, I perform a decomposition. As a whole the factor structure can account for 64 per cent of the gap and each factor plays an important role.

A key finding is that the parental valuation factor accounts for 22.5 per cent of the gender gap in participation. This is a much larger impact than one would find by restricting attention to observed variables such as parental aspirations. Parents influence their children's behavior in the teenaged years through several channels. The factor model makes it possible to quantify, and aggregate into the PVE factor, the different channels through which parents' valuations operate.

The rest of the paper proceeds by first describing how this paper relates to the existing literature. Next, I describe a simple model of the decision to enroll in university. The purpose of this model is guide the interpretation of the factors and to outline the conditions under which that interpretation is valid. Then, I describe the Youth in Transition Survey and how I use that data to estimate the unobserved factors affecting the gender gap in university enrollment. The results section commences with some reduced form regressions that describe key relationships in the data. The results from the factor model are then presented, followed by the decomposition of the university participation gap. The final part of the results section includes a discussion of why the parental valuation factor is higher among girls. Here, I address the issue of potential bias from unobserved ability. Finally, before concluding, I investigate whether the reasons parents give for their aspirations provide any clues as to why the PVE factor is, on average, higher for girls.

## 1.1 Related Research

Much of the economic literature investigating gender gaps in schooling and academic achievement is focused on explaining the trend. In this literature, the improving labour market conditions for women figure prominently (Goldin et al., 2006; Christofides et al., 2010; Fortin et al., 2015). Because I use data from a single cohort, this paper is more closely related to the literature examining cross-sectional gender differences in schooling outcomes. Specifically, I contribute Canadian evidence to the growing literature investigating how parents and family background function in the determination of gender gaps.

The female advantage in skills development that emerges in the early years is linked to parental behaviour and home environments. Using data from U.S., Canada, and the U.K., Baker and Milligan (2016) report that parents spend less time on teaching activities with boys and that difference helps explain why boys, ages four and five, have lower average scores on math and reading tests. In U.S. data, Bertrand and Pan (2013) document large gender differences in non-cognitive skills, specifically behavioural problems, in elementary school. Although Bertrand

and Pan (2013) conclude that parental time inputs do not help explain the non-cognitive skill gap, they do find that growing up in a single parent family increases the incidence of behavioural problems more for boys than for girls.

There is also further evidence from the U.S. that the negative consequences of growing up in less advantaged families continue to be more severe for males through secondary and post-secondary school and into adulthood. Autor et al. (2016), making use of linked administrative data from Florida, find that the gender gaps, which emerge in Kindergarten, are wider among those from lower socioeconomic backgrounds. That socioeconomic gradient in the gender gap is observed persistently throughout secondary schooling and high school graduation. Buchmann and DiPrete (2006) also find evidence from the National Educational Longitudinal Survey that females are more likely to complete college than males and that difference is largest in families where the father is either not present or has lower educational attainment. In data constructed from tax records, Chetty et al. (2016) show that the gender gap in employment rates which favours males who grew up in higher income families reverses among those in the bottom quintile of childhood family income.

Evidence drawn from Danish data presents a different pattern suggesting that the interaction between family background and gender gaps may vary across countries with differing institutional and economic contexts. Using Danish administrative data, Brenøe and Lundberg (2017) report that adolescent outcomes, such as completing grade nine on time, are better, on average, for girls. Like in the U.S., the gender gap in teenaged outcomes is exacerbated in less educated households. However, Brenøe and Lundberg (2017) go on to show that when considering educational attainment in adulthood, the pattern substantially changes, and also differs across mothers' and fathers' education levels. By age 27, the gender gap, which favours women, in years of schooling completed is wider for those whose mother has a university degree. In contrast, that gap narrows among those whose father has a university degree. Brenøe and Lundberg (2017) suggest that role model effects may in part explain why women may benefit relatively more from maternal education.

In my work, I am similarly interested in whether boys and girls are differentially impacted by their family background. Rather than focusing on socioeconomic characteristics such as income and parental education, I seek to identified a channel of parental influence related to their perceptions of the value of education. This paper is also among the few papers that provide evidence about gender gaps in schooling in Canada.

Investigations of gender differences in Canadian university enrollment, in particular, are constrained by a scarcity of suitable data. Card et al. (2011) make use of administrative data from the Ontario applications system to uncover any links between the gender gap in applications and in university enrollment. In 2006, there was a 13 percentage point difference, between young women and men, in the rates of application to Ontario universities. Although this data does not contain detailed information about family background, Card et al. (2011) perform a school level analysis linking the surrounding neighbourhood characteristics to the school. That analysis reveals that school characteristics explain little, and the gender gaps in application rates are similar whether a school is located in a low or high income neighbourhood.

Other Canadian evidence is drawn from the same data used in this paper–the Youth in Transition Survey (YITS). Frenette and Zeman (2007) decompose the gender gap in university attendance at age 19, using a Oaxaca-Blinder decomposition. They find that grades and performance on a standardized reading skills test are the most important explanatory variables. Study habits and parental aspirations are also significant contributors. Wage premiums are less important, explaining roughly 5 per cent of the gap. Christofides et al. (2008a,b) also use the YITS, attempting to relate the gender gap in youths' aspirations during high school to the gender in university participation. Their approach does not allow for any dependence between the unobserved components of youths' reported aspirations and university attendance.

My work extends Frenette and Zeman (2007) and Christofides et al. (2008a,b) by investigating the unobserved factors which drive the observed relationships. In particular, I distinguish between the stock of skills a 15 year old has accumulated and the value parents place on education. I describe the assumptions under which the impact of those factors are separately identified. While this paper confirms the importance of skills, and all of the parental investments in the early years helping to generate to those skills, the results also suggest that parents can have a direct impact on enrollment in the teenage years in a manner that helps explain the gender gap university attendance.

## 2 A simple model of university enrollment

In this section, I outline a simple economic model describing how the decision to enroll in university depends on three economic fundamentals, the earnings returns net of direct costs, non-pecuniary returns, and the rate at which youth can access funds for university. I use the model to specify how ability affects enrollment through each of these three channels, and to identify the sources of variation in parental valuations of education.

The model identifies three relevant stages in the enrollment decision, childhood (period 0), the teenage years (period 1), and young adulthood and beyond (period 2). This section begins with period 2, in which individuals, taking their ability and their parent's valuation of education as given, choose to attend university when doing so increases their utility. The ability that youth have accumulated by the time they decide to enroll in university evolves in the earlier periods. I describe that process of skills development, into which previous skills and parental investments are inputs. This section also defines the parental valuations of education and suggests some reasons in the context of the model why this factor might vary across males and females.

### 2.1 Period 2: Young adulthood and beyond

In period 2, young adults have the option to attend university or pursue an alternative path, which could involve enrolling in non-university courses or entering the labour force without further education. I assume that the decision to enroll in university depends linearly on three functions that capture the earnings returns to university, net of direct costs, the rate at which students can access funds during university, and non-pecuniary benefits from schooling. Respectively, these functions are labeled $g^W$, $g^R$, and $g^{NP}$.[3]

After accounting for direct costs, the university earnings premium is the difference between the present value of lifetime earnings from enrolling in university, and the opportunity cost of earnings from the next best alternative. I assume that these returns depend directly on one's ability, $\Theta$. To the extent that grades influence the type of program and the quality of university into which a youth is admitted, the university earnings premium also depends indirectly on high school grades, $grds_i$, in the following way:[4]

$$g_{it}^W = \alpha_i^W + \alpha_\Theta^W \Theta_t + \alpha_{grd}^W grds_i \tag{1}$$

I assume that there is no fundamental uncertainty in the model, meaning that individuals and their parents act as if they know the future returns to university. However, knowledge about future labour market returns vary across families, and parents influence what children understand about these returns. Heterogeneity in this knowledge will depend partly on the parents' own education level if, for example, more educated parents have access to better information about labour market returns (Junor and Usher, 2004):

$$\alpha_i^W = \alpha_{PE}^W PE_i + \eta_i^W \tag{2}$$

University enrollment also depends on the non-pecuniary returns to education. These non-financial costs and benefits include any direct consumption value or disutility of effort while in school. For this reason, $g^{NP}$ varies with a child's ability. Additionally, $g^{NP}$ reflects non-

[3] I show how this formulation relates directly to a life-cycle optimization problem in Appendix XXX.

[4] Because my goal is to focus on the student-side of the decision to pursue university, I do not explicitly model the admissions process. The relationship between high school grades and the future university earnings premium can be viewed as a linearization of that process. As a consequence of $grds_i$ entering (1) linearly, I am assuming that high school grades always increase earnings in the university sector more than in the alternative sector. This would be true if there is more variation in university quality and if university programs that lead to higher paying jobs are more selective than the alternative sector. Since grades do not directly impact the probability of attending university, I am implicitly assuming that students can always enroll in courses leading to a degree. By age 21, there are avenues for students with poor grades, or without a high school diploma to enroll in university courses through 'mature student' programs.

pecuniary returns to university that accrue over the life cycle, such as better health and higher quality marital matches (Oreopoulos and Salvanes, 2011). Parents can also impact the utility a child derives from attending university. This could occur if parents who value education offer emotional encouragement, and children enjoy earning their parents' approval. Additionally, as Bisin and Verdier (2001) suggest, parents may directly transmit a preference or taste for education to their children. With this in mind, I allow $g_{it}^{NP}$ to vary with parent's own education and unobserved preferences for education $\eta_i^{NP}$

$$g_{it}^{NP} = \alpha_{\Theta}^{NP}\Theta_t + \alpha_{PE}^{NP}PE + \eta_i^{NP} \tag{3}$$

Finally, among youth who face the same level of costs, and expect the same benefits, the decision to attend university may vary with a potential student's ability to access credit.[5] Following Cameron and Taber (2004), I assume that students can borrow at rates that differ during and after their schooling. The rate at which students can borrow while attending university depends on their parents in two ways. First, parental income and wealth influences whether children can access private or public student loans. Second, even holding resources constant, parents differ in their willingness to pay for their children's education. For example, some parents may feel that children will work harder if they have a 'horse in the race'. Other parents may use their resources to either induce children to attend or not to attend university. Parents' willingness to pay could also depend on how much they think a child will benefit from university, either in a pecuniary or non-pecuniary way. Therefore, assuming a linear form, $g_{it}^{R}$ is:

$$g_{it}^{R} = \alpha_0^R + \alpha_{PE}^R PE + \alpha_{FI}^R FI + \alpha_W^R\left(\alpha_{PE}^W PE_i + \eta_i^W\right) + \alpha_{NP}^R\left(\alpha_{PE}^{NP}PE + \eta_i^{NP}\right) + \eta_i^R \tag{4}$$

Family income ($FI$) and parental education are used to capture family wealth. The error term, $\eta_i^R$, in (4) reflects heterogeneity in the willingness to pay, as well as unobserved ability to pay.

[5]The literature.....

Adding $g^R$, $g^W$, and $g^{NP}$ together creates an index function for the decision to enroll in university, which is not yet a reduced form because $g^W$ is a function of high school grades. Grades are produced by a child's effort, their skills, and schooling and parental inputs.

$$grds_i = \alpha_0^G + \alpha_\Theta^G \Theta + \alpha_{PE}^G PE + \alpha_W^G \eta_i^W + \alpha_{NP}^G \eta_i^{NP} \tag{5}$$

Equation (5) is expressed as a reduced form evaluated at the youth's optimal effort level. Parental effort in helping their children can also impact grades (Houtenville and Conway, 2008). Thus, parental education enters (5) because their ability to help may depend on their own education level. Parent's might also encourage or enforce study effort and since this may depend on how parents view the returns to university $\eta_i^W$ and $\eta_i^{NP}$ enter equation(5).

After substituting equation (5) into $g^W$ and simplifying, a reduced form index function for university enrollment is:[6]

$$U_{it} = \gamma_0 + \gamma_{PE} PE + \gamma_{FI} FI + \Gamma_\Theta \Theta + \left(1 + \alpha_W^G + \alpha_W^R\right) \eta_i^W + \left(1 + \alpha_{NP}^G + \alpha_{NP}^R\right) \eta_i^{NP} + \eta_i^R \tag{6}$$

The reduced form coefficients reflect net effects on university enrollment from each variable's combined impact on perceived returns to university, both pecuniary and non-pecuniary, and the rate at which youth can finance their schooling. The last three terms in (6) reflect the unobserved variation in how parents' value education, which are the underlying source of variation in the PVE factor.[7]

Because the goal of this paper is to quantify the impact of parents' valuations on university enrollment beyond their impact on children's skills, the parental valuation factor is defined conditional on ability. Letting $\eta_i = \left(1 + \alpha_W^G + \alpha_W^R\right) \eta_i^W + \left(1 + \alpha_{NP}^G + \alpha_{NP}^R\right) \eta_i^{NP} + \eta_i^R$, the parental valuation factor in period $t$ is defined by a linear projection of $\eta_i$ on $\Theta_{it}$:

[6]Where, $\gamma_0 = \alpha_W^G \alpha_0^G + \alpha_R^0$, $\gamma_{PE} = \alpha_{PE}^W \left(1 + \alpha_W^R\right) + \alpha_{PE}^{NP} \left(1 + \alpha_{NP}^R\right) + \alpha_R^{PE} + \alpha_G^W \alpha_{PE}^G$, $\gamma_{FI} = \alpha_{FI}^R$, and $\Gamma_\Theta = \alpha_\Theta^W + \alpha_W^G \alpha_\Theta^G + \alpha_\Theta^{NP}$

[7]Although I suggest multiple sources of variation in parental valuations, I am not able to differentiate between these sources.

$$
\begin{aligned}
v_{pit} &= \eta_i - Proj\left[\eta_i|\Theta_{it}\right] \\
&= \eta_i - \delta\Theta_{it}
\end{aligned}
$$

Thus, by construction $v_{pit}$ is orthogonal to $\Theta_{it}$. [8] Given this definition, the impacts of parental valuations of education are not constant throughout childhood and adolescence, and depend on the evolution of ability in the first two periods.

## 2.2 Periods 0 and 1: Childhood and the teenage years

Youth take their current stock of skills as given when they decide to enroll in university, however, in their earlier years that ability is evolving. I assume that a child is born with an endowment of ability that depends partly on inherited ($\Theta_F$) and individual-specific ($\iota$) components:

$$\Theta_0 = f_0\left(\Theta_F, \iota\right) \tag{7}$$

Considerable evidence now suggests that there is more than one dimension to the ability and skills that matter in schooling and labour market outcomes (e.g. **?**Almlund et al., 2011; Heckman and Mosso, 2014). Following that literature, I assume that $\Theta$, in all time periods, has two dimensions labelled 'cognitive' and 'non-cognitive' skills.

Models of skills development advanced by Cunha and Heckman (2008) and Cunha et al. (2010) suggest that a child's future ability depends on their current level of ability, as well as the timing and level of parental investments. Investments in children's skills depend on family resources, including financial and time inputs, and the parents' own skills. Parental inputs might also vary according to their own understanding of how children acquire skills (Cunha et al., 2013). Based on those ideas, I assume that a child's ability evolves according to the following function:

$$\Theta_t = f_t\left(\Theta_{t-1}, I_t^p\left(PE, v_{pt-1}\right)\right) \tag{8}$$

[8] Any correlation between $\eta_i$ and $\Theta_{it}$ is captured by $\delta$ and in the empirical specification $\delta$ becomes a part of the factor loading on $\Theta_{it}$.

The second argument in (8) describes parental inputs as an investment function, $I_t^p\left(PE, \upsilon_{pt-1}\right)$, that depends on the parents' current valuation of education. Investments also depend on parents' education which reflects the amount and quality of time and financial inputs. Although, I do not explicitly model parents' investment decisions, I assume that they are choosing their investment level optimally, given the constraints on their resources and knowledge.

## 2.3 Interpreting the parental valuation of education

The dynamic interaction between the factors in (8) makes it clear that how one interprets the relationship between parental valuations and university enrollment depends critically on the time period in which $\Theta_t$ and $\upsilon_{pt}$ are evaluated. In the empirical model, the decision to attend university is expressed as a function of $\Theta_1$, ability at 15 years old. Because I assume that what matters is the level of ability, not how those skills are generated, conditioning on $\Theta_1$ is empirically equivalent to conditioning on all past values of ability and parental valuations. Moreover, if $\Theta_1$ is a sufficient statistic, then $\upsilon_{pi1}$ is also uncorrelated with the past value of $\Theta_0$.

Holding age-15 ability constant, the marginal effect of $\upsilon_{p1}$ on enrolment is not the total effect of parental valuations, because, over the years, all the influence that parents have had over their child's ability is held constant with $\Theta_1$. Instead, the impact of $\upsilon_{p1}$ is a contemporaneous impact, capturing influences that occur at age-15 and in the years leading up to the decision to attend university.

A further implication of the specification in (8) is that one of the channels by which $\upsilon_{p1}$ may affect university enrollment is by impacting ability in the years between age-15 and university enrollment.[9] Thus, in the same way that $\upsilon_{p1}$ would tend to under-estimate the total impact of parents' valuations, one would expect an even smaller impact of the PVE factor in the final period, (i.e. $\upsilon_{p2}$). Although the choice to evaluate ability and parental valuations at age-15 is partly data driven, this stage has particular policy relevance since it is the last phase of

[9] The evidence from Cunha and Heckman (2008) and Cunha et al. (2010) suggests that, in the adolescent years, changes along the non-cognitive dimension are more likely than along the cognitive.

compulsory schooling.

## 2.4 Gender differences in the factors

Until this point, I have described the decision to enrol in university without reference to gender. Differences in the distribution of any of the factors or in their impact might produce a gender gap in university enrollment.

There are a number of reasons to expect that 15 year-old boys and girls differ in their skills and abilities. While the role any biological differences might play in cognition is not well-understood, there is little doubt that a child's environment is an important factor, having direct, mediating, and moderating effects (Halpern, 2012). Thus, any differences in the types of families in which boys and girls tend to grow up might lead to gaps in their abilities (Bertrand and Pan, 2013). Parents might also invest differentially in boy's and girl's development, either because of differences in costs or the returns to investment (Baker and Milligan, 2016).

In accordance with its definition, gender differences in the PVE factor must be unrelated to skills but could be related to heterogeneity in how parents perceive the earnings and non-pecuniary returns, or their willingness to pay for university for equally skilled boys and girls. University wage premiums are generally larger for women (Boudarbat et al., 2010), even after controlling for ability (Caponi and Plesca, 2009). This is not because women with university degrees earn more than men, but because their outside options are relatively worse. In the 2006 Census, average earnings among young men with a certificate or diploma below the Bachelors level are roughly similar to that among women with a Bachelors degree (Turcotte, 2012). A gender gap in the PVE factor could arise directly from variation in parental perceptions of the differential returns. The differential returns to university might also indirectly shape parents' valuations through their willingness to pay. For example, parents may be less willing to pay for university, if they believe their sons can get a good job without a university degree, an option less available to daughters.

Parents might also evaluate the non-pecuniary returns to university differently by gender, and returns in the marriage market could be key among those differences. Schooling has value in the marriage market because it helps attract a desirable match, and because it can increase one's well-being within a marriage (Becker, 1973; Goldin, 1992; Peters and Siow, 2000). Chiappori et al. (2009) and Chiappori et al. (forthcoming) show that the returns to education in the marriage market can be higher for women when technological advancements reduce the time needed for home production, and as investment in children's human capital becomes more important. Echevarria and Merlo (1999) link gender differences in education to parental investments in an intergenerational household bargaining model. Altruistic parents make investments in their children's education, conditional on gender, taking into account the potential returns to education within a future marriage.

# 3 Data

The Youth in Transition Survey (YITS), a longitudinal survey of youth, is among the only data sets in Canada that combines information about academic achievement, attitudes and motivations, secondary and post-secondary schooling outcomes, and family background. Cohort A, which I use in this paper, is a nationally representative sample of Canadian youth born in 1984. The original sample, consisting of 29,687 students, was selected in two-stages. In the first stage, high schools were randomly selected from a list generated by the provinces. In the second stage, students were selected from within the schools to facilitate school-level analysis.[10] 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 I report, I use weights provided by Statistics Canada that account for over-sampling, as well as attrition.

In 2000, during the first cycle of the survey, students completed the Program for International Student Assessment (PISA) reading test. PISA tests, which are coordinated by the OECD, are

[10] 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.

designed to produce internationally comparable measures of knowledge and skills. A random subset, amounting to slightly more than half, of the students also wrote the math or science PISA tests. Because the sample sizes are so much smaller, I do not use those scores in the main estimation but make use of them in a robustness analysis.

The YITS also includes a parents' survey completed by the parent or guardian who identified him or herself as 'most knowledgeable' about the child. Parents provided information about themselves and their spouses, including their education and income. Parents also answered questions about their attitudes and behaviour as related to their children's education. The final component of the first cycle of YITS data collection is a school administrators survey, which collected information describing the schools' characteristics and resources.

Only the students were followed in the longitudinal component, and they were interviewed every two years. I combine data from the first cycle defining individual and family characteristics, with data from the fourth cycle, collected in 2006 when the students were 21 years of age. University participation is then defined as ever having enrolled in a program that leads to a Bachelors degree by age 21.

The analysis sample used in this paper is restricted to youth who completed the fourth survey, and whose parents completed the survey in cycle one. The final sample size among those with non-missing data is 7374 girls and 6805 boys. That there are more girls than boys in the sample is indicative of differential attrition. Although the weights do account for attrition based on observed characteristics, it is not possible to rule out the possibility of non-random attrition based on unobserved characteristics. Motte et al. (2008) provide additional information about the YITS and attrition.

For the pooled sample, and for boys and girls separately, in Table 1, I report the variables means. The definition of the variables are provided as they are introduced in the paper.

# 4 Empirical Approach

There are a number of ways in which one might approach estimating the reduced-form for university enrollment derived in Section 2. I employ a factor model that was developed in FGG, and which is an extension of the approach developed by Carneiro et al. (2003) and Cunha et al. (2005), hereafter CHH and CHN. In this section, I discuss the factor model and the advantages it has over other approaches. To begin, it is useful to write out the linear index function for university enrollment, that follows from equation (6), after substituting in the parental valuation of education, $v_{ip}$, evaluated at age-15:[11]

$$U_i^g = \gamma_0 + \gamma_f female_i + \gamma_z Z_i + \lambda^g_{0\theta_1}\theta_{i1} + \lambda^g_{0\theta_2}\theta_{i2} + \lambda^g_{0v_p} v_{ip} + u^g_{i0} \tag{9}$$

In equation (9), $Z_i$ is a vector of observed characteristics, including parents' education, and family income. The two dimensions of $\Theta$, $\theta_1$ and $\theta_2$, are cognitive and non-cognitive ability, respectively. The intercept shifter, $\gamma_f$, captures any difference between males and females in the propensity to enroll in university, holding the observed and the unobserved characteristics constant.

If one ignores the unobserved factors, and estimates (9) by OLS or in a simple Probit, the estimator for $\gamma_f$ will be biased by any of the gender differences in unobserved factors discussed in Section 2.4. A common approach to control for unobserved factors is to use observed variables to stand in as proxies. This is essentially the approach taken by Frenette and Zeman (2007).

If proxies imperfectly measure the unobserved phenomenon of interest, that measurement error can lead to biased estimates. If there is a second measure of the unobserved factor, it can be used as an instrument for the proxy (Griliches and Mason, 1972). In the simple case with a single unobserved factor, one outcome, and two measurements, the instrumental variables (IV) estimator for the coefficient on the unobserved variable is equivalent to the factor model. I demonstrate this in an appendix.

[11]In this section, I suppress the time subscripts. The independent variables are all evaluated at 15 years of age.

The factor model has an important advantage over the IV estimator, which is particularly useful in understanding the gender gap in university attendance. The IV estimator only accounts for shifts in the mean of the unobserved factors. In contrast, the factor model estimates a *distribution* for each unobserved characteristic. I allow the factor structure to vary across males and females, which means the model is flexible enough to capture any non-linearities in the relationship between gender and unobserved variables. The results in Fortin et al. (2015) suggest that gender gaps in student achievement are much larger at the top of the distribution. If gender does affect the shape of the factor distributions then ignoring those non-linearies will bias the estimated gap in enrollment.

Following the approach outlined in CHH, I estimate the distributions of the unobserved variables by imposing covariance restrictions on a system of noisy measures of the factors. One of the conditions for identification is that there are at least two measurements for each unobserved factor.

The first factor, $\theta_{i1}$, is measured using quartiles of the PISA reading score. The index for those PISA quartiles is[12]:

$$PISA_i^g = \beta_{10}^g + \beta_{1z} Z_i + \lambda_{1\theta_1}^g \theta_{i1} + u_{i1}^g \quad (10)$$

This specification, which implies that the factor loads on both $\theta_{i2}$, and $\upsilon_{ip}$ are zero in the PISA measurement equation, is an important restriction and affects the way the factors are interpreted. Drawing from the model developed in Section 2, $\Theta$, represents the stocks of skills accumulated up to age 15. The first element, $\theta_{i1}$, is labeled the 'cognitive skill' factor.

There is evidence, however, that performance on low stakes tests depends on personality characteristics that differ from what might be typically thought of as 'cognitive skills' (Borghans et al., 2011). For example, low test scores may reflect a lack of motivation rather than ability.

[12] I use the average across the 5 plausible values for the Reading test. Then, I generate quartiles across the full YITS sample, including non-responders to the fourth cycle and the parental survey. In other words, the quartiles are defined before any sample restrictions are made.

With this in mind, $\theta_{i1}$ encompasses cognitive ability and any traits associated with test-taking effort, such as the 'desire to please'.

The second factor, $\theta_{i2}$, is labelled as 'non-cognitive' skills, and because the factors are orthogonal by construction, it will reflect characteristics that are not already captured by $\theta_{i1}$. Although I do not cleanly and separately identify cognitive from non-cognitive skills, interpretation of the third factor does not require that. Instead, it is important that together $\theta_{i1}$ and $\theta_{i2}$ represent skills as completely as possible, such that when they are held constant, the impact of $\upsilon_{ip}$ can be interpreted as an effect of parents during the teenaged years.

The exclusion of $\upsilon_{ip}$ from equation (10) is another important restriction. Because PISA is a one-time, low-stakes test, parents are unlikely to directly influence the test score by, for example, helping their children prepare. The PISA test is not used in assessing individual student performance, nor does it measure mastery of course curriculum. Neither parents nor the students in the YITS obtained their individual PISA test scores. For these reasons, $\upsilon_{ip}$ does not enter the PISA equation.

For the second factor, $\theta_{i2}$, the measures are selected to best reflect the concept of 'Conscientiousness', which is a part of the 'Big Five' taxonomy of personality characteristics. This personalty trait is characterized by the adjectives: efficient, organized, planful, reliable, responsible, and thorough (McCrae and John, 1992). Measures of conscientiousness have also been found to predict educational outcomes, including grades in post-secondary schooling and years of education (Borghans et al., 2008; Almlund et al., 2011).

Although the YITS does not contain a specific scale for conscientiousness, following FGG, I use measures of self-reported behaviours that are related to the characteristic adjectives. The first measure is a variable that takes on the value one if the youth responded 'always' when asked how often the following statement applies: 'I complete my homework on time'. Homework is not always graded, as such its timely completion is at least to some extent 'voluntary' and reflects a level of conscientiousness.[13] The underlying index function is:

$$hmwrk_i^g \quad = \quad \beta_{20}^g + \beta_{2z} Z_i + \lambda_{2\theta_2}^g \theta_{i2} + \lambda_{2v_p}^g v_{ip} + u_{i2}^g \tag{11}$$

In FGG, and Heckman et al. (2006), the non-cognitive measures are not a function of the cognitive factor. I follow suit here. However, even after controlling for non-cognitive skills, parents' can influence the timely completion of homework by offering incentives or punishments, and as such the parental valuation factor is included in the *hmwrk* equation.

The key measurement for the parental valuation of education is the parental aspirations question. The responding parent was asked 'What is the highest level of education that you hope your child will get?'. The corresponding variable is coded as equalling one if the parent responded either 'One university degree' or 'More than one university degree'. Because parents almost certainly take into account not only their own valuation of education but also their children's current stock of skills, the parental aspiration measurement is a function of all three factors:

$$parasp_i^g \quad = \quad \beta_{30}^g + \beta_{3z} Z_i + \lambda_{3\theta_1}^g \theta_{i1} + \lambda_{3\theta_2}^g \theta_{i2} + \lambda_{4v_p}^g v_{ip} + u_{i3}^g \tag{12}$$

Since identification requires at least two measurements for each factor, for another measure of cognitive ability, I use overall high school grades reported by the youth at age 15. Grades generally reflect, not just academic skill, but also effort and behaviour in school. Moreover, since parents can become involved in their child's school work either directly or though encouragement, grades vary with all three factors[14]:

$$grades_i^g \quad = \quad \beta_{40}^g + \beta_{4z} Z_i + \lambda_{4\theta_1}^g \theta_{i1} + \lambda_{4\theta_2}^g \theta_{i2} + \lambda_{4v_p}^g v_{ip} + u_{i4}^g \tag{13}$$

[13]Eight percent of girls and 11 percent of boys report that their teacher always mark their homework.

[14]In FGG, to help justify the inclusion of the parental valuation factor in the *grades* equation while it is excluded in the PISA reading score equation, we show that Math and Sciences grades are significantly related to both reading scores and parental aspirations. In contrast, PISA math and science scores, after controlling for the reading scores, are not related to parental aspirations.

The second measure of non-cognitive skills is related to the 'thoroughness' aspect of Conscientiousness. This variable takes on the value one if the youth responded 'never' when asked how often the following statement was true, "I do as little work as possible; I just want to get by." The underlying index function is:

$$getby_i^g = \beta_{50}^g + \beta_{5z} Z_i + \lambda_{5\theta_2}^g \theta_{i2} + \lambda_{5v_p}^g v_{ip} + u_{i5}^g \quad (14)$$

I use a variable that indicates whether parents have saved for their children's education as a second measure of parental valuations. Specifically, parents are first asked "Have you (or your partner) done anything specific to ensure that your child will have money for further education after high school?" Having 'saved' means the parent further indicated that he or she had 'started a savings account', 'started a Registered Education Savings Plan (RESP)', 'set up a trust fund for this child', or 'made investments, such as mutual funds or Canada Savings Bonds'. Like the parental aspirations variable, the 'saved' index is a function of all three factors:

$$saved_i^g = \beta_6^g + \beta_{7z} Z_i + \lambda_{6\theta_1}^g \theta_{i1} + \lambda_{6\theta_2}^g \theta_{i2} + \lambda_{6v_p}^g v_{ip} + u_{i6}^g \quad (15)$$

Finally, I include a seventh variable that measures youth's own aspirations and which over identifies the model. Youth were asked at age 15 'What is the highest level of education you would like to get?'. The index function for youths' aspirations is:

$$yasp_i^g = \beta_{70}^g + \beta_{7z} Z_i + \lambda_{7\theta_1}^g \theta_{i1} + \lambda_{7\theta_2}^g \theta_{i2} + \lambda_{7v_p}^g v_{ip} + u_{i7}^g \quad (16)$$

The sample variances and covariances between the outcomes and the measurements provide the identifying information. The distributions of the unobserved factors are not however recovered without further normalization and assumptions. In particular, the factors are each mutually independent, with a mean of zero. The measurement errors ($u^g$) are also independent of the covariates, the factors and other errors. At a minimum, the number of measurements should be twice the number of factors, plus one.

Since the factors have no natural scale, for each factor, one of the loadings is normalized to one. Here, the cognitive factor loading in the $PISA$ measurement is normalized, as are the non-cognitive and PVE loadings in the $hmwrk$ and $parasp$ equations, respectively.

Finally, the restriction that there is one measurement which is dedicated to a single factor is necessary for identification.[15] That dedicated measure is the PISA equation which is a function of only the cognitive factor. Carneiro et al. (2003) prove that unobserved factor distributions are identified under these conditions, and I provide more details of how the factors are identified in this model in an Appendix.

I estimate two versions of the model. The first is a 'constrained' model in which the parameters are the same for men and women, with the exception of the female intercept in each equation. The 'flexible' or 'unconstrained' model, allows the factor structure to differ across genders. Similar to the approach taken in Heckman and Singer (1984), the factors are specified as discrete variables, where one point of support is normalized to zero. Since these factors have no meaningful scale, the factor locations are the same for both sexes in the flexible model, however, the probability associated with each level depends on gender. Additionally, the factor loadings, or the impact the factors have on university participation, differ by gender.[16]

Both versions of the model are estimated by maximum likelihood. The likelihood function is defined conditional on each level of the factors, then weighted by the probability associated with each factor location and summed. An example contribution to the likelihood function, in the flexible model, is:

[15] This normalization is described in footnote 18 of Carneiro et al. (2003).

[16] The coefficients on the observed variables are constrained to be the same for both genders in all of the equations. Allowing these to vary in a model where the factor locations are also the same makes the model computationally infeasible. I have, however, also estimated the model separately by gender, which allows every parameter to differ. With these models it is not possible to compare the scale of the factors. However, I can gauge whether differences in the coefficients on the observed variables contribute substantially to the gender gap. Generally, the role they play is relatively small and the rest of the results change little. These results are available by request.

$$\begin{aligned}\sum_{\theta_1}\sum_{\theta_2}\sum_{\upsilon_p} p^g(\theta_1)\,p^g(\theta_2)\,p^g(\upsilon_p)\quad & F\left(\gamma_0^g+\gamma_z Z_i+\lambda_{0\theta_1}^g\theta_{i1}+\lambda_{0\theta_2}^g\theta_{i2}+\lambda_{0\upsilon_p}^g\upsilon_{ip}\right)*\\ & F\left(\sigma_p^{-1}\left[PISA_1-\beta_{10}^g-\beta_{1w}Z_i-\theta_{i1}\right]\right)*\\ & F\left(\beta_{20}^g+\beta_{2w}Z_i+\theta_{i2}+\lambda_{2\upsilon_p}^g\upsilon_{ip}\right)*\\ & F\left(\beta_{30}^g+\beta_{3z}Z_i+\lambda_{3\theta_1}^g\theta_{i1}+\lambda_{3\theta_2}^g\theta_{i2}+\upsilon_{ip}\right)*\\ & \sigma_g^{-1}f\left(\sigma_g^{-1}\left[grds_i-\beta_{40}^g-\beta_{4w}Z_i-\lambda_{4\theta_1}^g\theta_{i1}-\lambda_{4\theta_2}^g\theta_{i2}-\lambda_{4\upsilon_p}^g\upsilon_{ip}\right]\right)*\\ & F\left(\beta_{50}^g+\beta_{5w}Z_i+\lambda_{5\theta_2}^g\theta_{i2}+\lambda_{5\upsilon_p}^g\upsilon_{ip}\right)*\\ & F\left(\beta_{60}^g+\beta_{6z}Z_i+\lambda_{6\theta_1}^g\theta_{i1}+\lambda_{6\theta_2}^g\theta_{i2}+\lambda_{6\upsilon_p}^g\upsilon_{ip}\right)\\ & F\left(\beta_{70}^g+\beta_{7z}Z_i+\lambda_{7\theta_1}^g\theta_{i1}+\lambda_{7\theta_2}^g\theta_{i2}+\lambda_{7\upsilon_p}^g\upsilon_{ip}\right)*\end{aligned}\tag{17}$$

This example contribution is for an individual, of gender $g$, who attended university, had university aspirations, scored in the bottom PISA quartile, always turned in their homework on time, and never just wanted to get by. This individual's parent also had university level aspirations for their child and saved for their child's education.

The $F()$'s are cumulative normal distribution functions; $f()$ is the normal pdf; the $p(.)$'s are probabilities associated with the points of support; and, $Z$ is a vector of observed variables which includes, parental education, family income, indicators for rural residence, immigrant status, and living in a two-parent family. This vector also includes province dummies and distance from the students' high school to the nearest university. The $W$ vector includes all the variables in $Z$ except the distance to nearest university, and it enters equations where the costs of university are unlikely to directly affect the outcome.

With the exception of the $PISA$ and $grades$ equations, each component of the likelihood is a Probit. Grades enter linearly and the PISA quartiles are modeled as an ordered Probit, where $PISA_1$ is the cut-off value between the first and second quartile. The standard deviations of PISA and grades are $\sigma_p$ and $\sigma_g$, respectively.

# 5 Results

Before discussing the factor model, I begin with results from Probit regressions where the measurement variables act as proxies for the underlying factors. Although the proxy approach to estimating the gender gap is inconsistent, these simple models help describe key patterns and correlations in the data. The marginal effects from these regressions are presented in Table 2.[17] The dependent variable in each case is the dummy variable that equals one if the youth attended university by age 21 and zero otherwise.

The first regression includes only socio-economic and household characteristics. These characteristics include a set of six dummy variables describing the parents' highest level of education. The reference category is 'both parents did not finish high school'. The remaining categories are: Both parents have a Bachelors degree or higher, one parent has a Bachelors or higher, both parents have a post-secondary education (PSE) credential that is not a university degree, only one parent has PSE below the Bachelors level, both parents have a high school diploma, and one parent has a high school diploma. Lone parent families are coded into the 'both parents' categories.

The set of socioeconomic variables also includes indicators for whether the youth lives in a two parent family, in a rural area, is an immigrant, or is Indigenous. The natural log of family income is also included. I adjust family income to account for household size by dividing total before-tax family income by the square root of the number of household members. All of these family background characteristics are reported by the parents when the youth was age 15.

Parental education is the strongest predictor of university participation, a result that is commonly found in this and other Canadian data sets (Drolet, 2005; Christofides et al., 2009; Finnie and Wismer, 2011). The remaining socioeconomic variables have the expected effect on university participation. The gender gap after controlling only for socioeconomic variables is 15

[17]The coefficients from these models as well as models estimated separately for males and females are reported in an appendix.

percentage points, which is roughly the same as the unconditional gap.

The next three regressions, reported in columns (2) through (4), introduce the sets of measurement variables separately, then the fifth regression includes all the measurements. In these models, the measurements act as proxies for the underlying factors. The gender gap narrows in each of these specifications, particularly when all the proxies are entered simultaneously. In that regression, females are only 5 percentage points more likely to attend university. Moreover, each of the measurements significantly predicts enrollment. The indicator for whether a youth "just wants to get by" is insignificant in the specification with all proxies because of its correlation with grades.

In the final regressions, I include school and peer characteristics. The school characteristics include an index of school quality reported by the high school administrator, the ratio of students to teachers, and the ratio of boys to girls. None of these are statistically significant. There is evidence from U.S. data that students are more favourably evaluated by teachers who are similar in terms of gender and race (Dee, 2005). Unfortunately, the YITS does not contain specific information about the characteristics of each student's teachers.

There is also evidence that one's peers, and in particular the gender of those peers, affects achievement in high school (Hoxby, 2000; Hill, 2015). Although the type of information needed to identify peer effects is not available in the YITS, the youth are asked about their closest friends. I include an indicator for whether the youth said that all of their friends 'think completing high school is very important?'. This variable is statistically significant but the effect size is fairly small.

Overall, the strength of association between enrollment and the measurement variables can be contrasted with that of the school and friends' characteristics. Moreover, the variables included in the final regression have little impact on the estimated marginal effect of the female indicator.

## 5.1 Factor Model

In this section, I discuss the results from the factor model introduced in Section 4, beginning with the support of the factor distributions. The number of points of support for each factor was determined empirically. The model I present here has three points of support for the cognitive and non-cognitive factors and two points of support for the PVE factor. I began with two points of support and added a third point to each factor in turn. Both the Akaike Information Criterion and the Baysian Information Criterion rejected the model with three cognitive, three PVE and two non-cognitive points of support in favour of the model presented here.[18]

The parameter estimates from the university outcome equation are reported in Table 3, with the constrained factor model presented beside the flexible model, in which the factor distributions and loadings vary by sex. Comparing the two models, the coefficients on the observable variables in the university equation are quite similar, except for the female dummy, which shifts the enrollment intercept. Relative to the constrained model, the female dummy is roughly 30 % smaller and is no longer statistically significant. As I will discuss later, this is the first indication that the factors play a substantial role in explaining the gender gap in enrollment.

For the measurement equations, I report the intercepts and factor loads in Table 4. At the bottom of this table, the sum of the log likelihoods for each model are shown. Using these to construct a likelihood ratio test, I can reject the null hypothesis that the factor loadings and distributions are the same for boys and girls with a very high level of significance.[19]

For each of the factors, the loading parameters in the university outcome equation are statistically significant in both models. The size of the factor loadings are difficult to interpret on their own because the scale of any factor is determined by the scale of the measurement equation in which the load is normalized. Nonetheless, the statistical significance implies that each factor plays a role in determining university attendance. While the PVE factor is related to parental

[18] The model with three points of support in each factor, which is 27 different intercepts, never converged.

[19] The likelihood ratio test statistic is 442.74. The flexible model has 22 more parameters than the constrained model.

aspirations by construction, the cognitive and non-cognitive factors are also significantly related to parental aspirations. Taken together, these results imply that the factors represent possible channels through which parents' aspirations impact university participation.

To investigate the size of the factors' impact on university participation, for both genders, I plot a predicted probability evaluated for 12 of the different factor levels in Figure 3. I omit the lowest category of non-cognitive skills because only 6% of girls and 11% of boys fall into this category.[20] Within each smaller graph in Figure 3, the cognitive gradient is observed. For example, in the bottom right hand graph, for boys with a high level of all three factors the predicted probability of attending university is .9251. The likelihood of attending university among girls with high PVE and non-cognitive factors but a low cognitive factor is .3652. The comparison of these probabilities reveals a cognitive gradient of .5599. Within each row, the different graphs reflect levels of the non-cognitive factor and comparing the bottom and top rows reveals the impact of the PVE factor. I also report each marginal effect in Tables 5 to 7.

Before comparing across genders, there are several general observations that can be made about the predicted probabilities. First, all three factors have a large impact on university participation. Except among those with low non-cognitive skills and parental valuations, the cognitive skill gradients range from .44 to .63. The effect that the non-cognitive factor has on university enrollment is not quite as large. However, since the cognitive factor also captures skills that are associated with test-taking effort, the non-cognitive effects reported in Figure 3 and Tables 6 can be thought of as a lower bound.

Because the model is non-linear, the size of the non-cognitive impacts vary considerably depending on the level of the PVE factor at which the effect is evaluated. Taking as an example the probabilities predicted at the high cognitive level, when the PVE factor is low, the impact of having high non-cognitive skills–relative to the medium level– is 17.8 and 15.9 percentage points for boys and girls, respectively. In contrast, the same comparison when the PVE factor

[20] A figure with all 18 levels is reported in an appendix.

is high is 8.3 for boys and 7.3 percentage points for girls. Part of the reason for this is that the probability is already quite high when the PVE factor is high, as such there is not as much scope for non-cognitive skills to improve the chances of attending university.

Similarly, the parental valuation factor's effects also depend on the levels of the skills factors. The impact of the PVE factor tends to be largest at the medium level of the non-cognitive factor, ranging between .19 and .41. One way to understand the magnitude of the PVE factor effect is to compare the predicted probabilities in Figure 3 to the unconditional probabilities, which are .37 for boys, and .52 for girls. For girls, if their parent has a low valuation, their probability of attending university is at or below average unless they have both high cognitive and non-cognitive skills. Boys with a low PVE attend university with at least an average probability when they have high cognitive skills and at least the medium level of non-cognitive skills.

Although the impacts vary, the PVE factor increases the probability of attending university by a large margin for all the levels of skills [21] For children with high levels of both skills, having a parent who values education raises the chance of attending university from less than .7 to .93. This is very different from what we found in FGG when we studied the impact of parental valuations on the high school dropout decision. There, we found that the PVE had essentially no effect on children with high cognitive skills. This is because the vast majority of students in the YITS finish high school and having high cognitive skills is enough to virtually guarantee completion.[22] In contrast, taking boys and girls together fewer than half attend university (45%). For the university enrollment decision, while having high levels of cognitive and non-cognitive skills increases one's chances, there is still plenty of room for factors such as parents' valuation of education to make a difference in the teenage years.

Turning now to a comparison across genders, the most striking result is how much smaller

[21] The impact of the PVE factor is smaller when evaluated at the lowest level of both cognitive and non-cognitive skills, however, these skill levels are predicted to occur with low probability. Specifically, 0.012 for boys and 0.005 for girls.

[22] It is worth pointing out that the YITS sampling strategy did not include high schools where one might expect to find very high levels of socioeconomic disadvantage such as schools on First Nations reserves and in the Territories.

the gender differences are, once one conditions on the three factors. Among those with high cognitive skills, there is virtually no difference between boys and girls in the probability of attending university, no matter the level of the other factors. The largest gender gaps which persist are among youth with the low level of cognitive skills and high parental valuations. Girls are about 12 percentage points more likely to attend university than boys if they have low cognitive skills and high levels of non-cognitive and PVE factors. That same difference is 11 percentage points for the medium level of non-cognitive skills. Although these are quite large differences, which compare to the unconditional gap of .15, a relatively small fraction of the mass is estimated at those factor levels. Across all three levels of non-cognitive skills, roughly 9 per cent of both boys and girls have low cognitive skills and a high PVE. Additionally, because these predicted probabilities are the average of a non-linear function evaluated at a particular vector, the difference in any two probabilities stems partly from the concavity of the cumulative normal function. The differences will be more pronounced at lower levels where the function is more convex.

Since, after conditioning on the factors, the predicted probabilities of attending university are much closer, this implies different distributions of the factors for boys and girls. Figure 2 reveals that this is indeed the case. The top panel of Figure 2 shows the marginal distributions for each of the three factors. The bottom panel reports the joint distributions, which are the product of the marginal distributions because the factors are orthogonal.

Girls have higher average levels of each factor. They are more likely to have the higher level and less likely to have the lower level of cognitive skills. The predicted probability of having the highest level of non-cognitive skills is almost 8 percentage points higher for girls. Girls are also 10 percentage points more likely to have parents with the highest valuation.

The estimated distributions of the factors reported in Figure 2 also make it clear that gender differences in skills affect the whole distribution rather than just a mean difference. The female advantage is larger at the top of both skills distributions. An instrumental variables or proxy variable approach would not capture such distributional shifts.

## 5.2 How much of the gap is explained by the factors?

To further explore the role each factor and factor loading plays in explaining the gender gap in participation, I perform a decomposition exercise, following in the spirit of a Oaxaca-Blinder decomposition. I begin by expressing the unconditional gender gap in terms of the factor model. Unlike with a linear decomposition, it is important to take the average predicted probabilities, rather than evaluate the probability at the average (Fairlie, 1999; Fortin et al., 2011). If $F$ is the cumulative normal distribution, using the enrollment index in equation (**??**), the total gender gap is:

$$\begin{aligned}\Delta_U &= U\left(X^f, f\right) - U\left(X^m, m\right) \qquad (18)\\ &= \frac{1}{n^f}\sum_{i=1}^{n^f}\sum_{\theta_1}\sum_{\theta_2}\sum_{v_p} p^f\left(\theta_1\right)p^f\left(\theta_2\right)p^f\left(v_p\right)F\left(\gamma_0^f + \gamma_x X_i^f + \lambda_{0\theta_1}^f\theta_{i1} + \lambda_{0\theta_2}^f\theta_{i2} + \lambda_{0v_p}^f v_{ip}\right)\\ &- \frac{1}{n^m}\sum_{i=1}^{n^m}\sum_{\theta_1}\sum_{\theta_2}\sum_{v_p} p^m\left(\theta_1\right)p^m\left(\theta_2\right)p^m\left(v_p\right)F\left(\gamma_0^m + \gamma_x X_i^m + \lambda_{0\theta_1}^m\theta_{i1} + \lambda_{0\theta_2}^m\theta_{i2} + \lambda_{0v_p}^m v_{ip}\right)\end{aligned}$$

The model estimates the total gap as.1508. To calculate how much of that raw difference can be attributed to differences in the observed characteristics, I need to evaluate the predicted probabilities using the parameters from one gender and the $X$-vector from the other gender. Specifically:

$$U\left(X^g, h\right) = \frac{1}{n^g}\sum_{i=1}^{n^g}\sum_{\theta_1}\sum_{\theta_2}\sum_{v_p} p^h\left(\theta_1\right)p^h\left(\theta_2\right)p^h\left(v_p\right)F\left(\gamma_0^h + \gamma_x X_i^g + \lambda_{0\theta_1}^h\theta_{i1} + \lambda_{0\theta_2}^h\theta_{i2} + \lambda_{0v_p}^h v_{ip}\right)$$

Putting these terms together, yields the explained and unexplained differences from the familiar Oaxaca-Blinder decomposition, which can be implemented in two different ways[23]:

[23]A third decomposition is possible using parameters from the constrained or 'pooled' model.

Using the male parameters:

$$\Delta_{X^m} = \underbrace{U\left(X^f, m\right) - U\left(X^m, m\right)}_{\text{Explained}} + \underbrace{U\left(X^f, f\right) - U\left(X^f, m\right)}_{\text{Unexplained}} \quad (19)$$

Using the female parameters:

$$\Delta_{X^f} = \underbrace{U\left(X^f, f\right) - U\left(X^m, f\right)}_{\text{Explained}} + \underbrace{U\left(X^m, f\right) - U\left(X^m, m\right)}_{\text{Unexplained}} \quad (20)$$

The first term, in both (19) and (20), represents the part of the gender gap that can be explained by differences in the observed characteristics, and the second term represents the gap that is 'unexplained'. This decomposition is presented in the first row of Table 8.[24] The set of socio-economic variables included in the university enrollment index explain essentially none of the total gender gap. This conclusion does not depend on which set of parameters is used in the decomposition. This result is seemingly quite different from Frenette and Zeman (2007) who conclude that socioeconomic characteristics explain about three quarters of the gap. However, their set of socioeconomic characteristics includes variables such as parental aspirations and grades. The goal here is to try to disentangle the different factors that are reflected in those variables. The same covariates that I include in the university enrollment equations, such as parental education and income, similarly explain very little of the gender gap in Frenette and Zeman (2007).

Because the flexible factor model allows the distribution of factors and their loadings to vary by gender, I can also perform a decomposition based on the factors. In that case, the counterfactual is:

[24] The standard errors are obtained by bootstrapping. To account for the school based sampling frame of the YITS, the bootstrap sampling is performed at the school level. In the pooled sample there are 1098 schools and, respectively, in the male and female samples there are 1066 and 1055 schools. In each of 200 repetitions, I randomly draw a sample of schools from the data and a vector of parameters from the estimated sampling distribution.

$$U\left(X^g,\gamma_0^g,h\right) \;=\; \frac{1}{n^g}\sum_{i=1}^{n^g}\sum_{\theta_1}\sum_{\theta_2}\sum_{\upsilon_p} p^h\left(\theta_1\right)p^h\left(\theta_2\right)p^h\left(\upsilon_p\right)F\left(\gamma_0^g+\gamma_x X_i^g+\lambda_{0\theta_1}^h\theta_{i1}+\lambda_{0\theta_2}^h\theta_{i2}+\lambda_{0\upsilon_p}^h\upsilon_{ip}\right)$$

The fraction of the gap attributable to the factor structure is given by this decomposition:

Using the male parameters: (21)

$$\Delta_{\Lambda\Theta^m} = \underbrace{U\left(X^f,f\right)-U\left(X^f,\gamma_0^f,m\right)}_{\text{Explained by factor structure}} + \underbrace{U\left(X^f,\gamma_0^f,m\right)-U\left(X^m,m\right)}_{\text{Unexplained}}$$

Using the female parameters: (22)

$$\Delta_{\Lambda\Theta^f} = \underbrace{U\left(X^m,\gamma_0^m,f\right)-U\left(X^m,m\right)}_{\text{Explained by factor structure}} + \underbrace{U\left(X^f,f\right)-U\left(X^m,\gamma_0^m,f\right)}_{\text{Unexplained}}$$

Here, the unexplained portion is driven by the female intercept, and the very small differences in the $X$ vector. The portion explained by the factor structure can be further decomposed into components explained by each factor, or the factor loading. For example, to learn how much of the gap occurs because girls are more likely to have a higher level of the cognitive factor, I can construct the following counterfactual predicted probability:

$$U\left(g,p^h\left(\theta_1\right)\right) \;=\; \frac{1}{n^g}\sum_{i=1}^{n^g}\sum_{\theta_1}\sum_{\theta_2}\sum_{\upsilon_p} p^h\left(\theta_1\right)p^g\left(\theta_2\right)p^g\left(\upsilon_p\right)F\left(\gamma_0^g+\gamma_x X_i^g+\lambda_{0\theta_1}^g\theta_{i1}+\lambda_{0\theta_2}^g\theta_{i2}+\lambda_{0\upsilon_p}^g\upsilon_{ip}\right)$$

Then, I can calculate the part of the gap attributed to the distribution of $\theta_1$ with:

Using the male parameters: (23)

$$U\left(X^f,f\right)-U\left(X^f,\gamma_0^f,m\right) = \underbrace{U\left(X^f,f\right)-U\left(f,p^m\left(\theta_1\right)\right)}_{\text{Explained by } p^m(\theta_1)} + \underbrace{U\left(f,p^m\left(\theta_1\right)\right)-U\left(X^f,\gamma_0^f,m\right)}_{\text{Unexplained}}$$

Using the female parameters:

$$U\left(X^m,\gamma_0^m,f\right)-U\left(X^m,m\right) = \underbrace{U\left(m,p^f\left(\theta_1\right)\right)-U\left(X^m,m\right)}_{\text{Explained by } p^f(\theta_1)} + \underbrace{U\left(X^m,\gamma_0^m,f\right)-U\left(m,p^f\left(\theta_1\right)\right)}_{\text{Unexplained}}$$

Performing a detailed decomposition, that is assigning a portion of the gap to each factor and its loading, poses some problems because the model is non-linear.[25] In particular, if performed sequentially, the order in which I perform the decomposition can affect the results. Alternatively, if the decomposition is conducted piecewise, the components of the decomposition will not necessarily add up to the whole. Because the second approach is arguably more transparent, I have chosen to perform the decomposition by switching one factor or one loading at a time. The equations describing each decomposition, which take an analogous form to (23), are presented in an appendix.

Performing a detailed decomposition, which is assigning a portion of the gap to each factor and its loading, poses some problems because the model is non-linear.[26] In particular, if performed sequentially, the order in which I perform the decomposition can affect the results. Alternatively, if the decomposition is conducted piecewise, the components of the decomposition will not necessarily add up to the whole. Because the second approach is arguably more transparent, I have chosen to perform the decomposition by switching one factor or one loading at a time. The equations describing each decomposition, which take an analogous form to (23), are presented in an appendix.

As was the case with the observed characteristics, the factor decompositions, also reported in Table 8, are very similar whether I use the female or male coefficients to build the counterfactual. Using the male parameters, the factor structure, taken as a whole, accounts for .096 of the total gap of .1508, or 64 percent. I begin the detailed decompositions with the factor loadings, which, overall, are not as important as the distributions of the factors. The factor loading decompositions that are reported in Rows (3) through (5) indicate how much of the gap is generated by differences in the impact of the factors on the enrollment index. The loading on the cognitive skills factor explains 3 percentage points of the gap. The contributions of the other two two factor loadings are negative but not statistically significant.

[25] See Fortin et al. (2011) for a detailed discussion of the issues associated with non-linear decompositions.
[26] See Fortin et al. (2011) for a detailed discussion of the issues associated with non-linear decompositions.

Rows (7) through (10) of Table 8 report the part of the gap driven by the factor distributions. Each of the factors plays an important role. The cognitive skills factor accounts for 3.1 and 3.4 percentage points using the male and female parameters, respectively. The part of the gap attributable to non-cognitive skills is slightly smaller at 2.8 percentage points. Finally, the parental valuation of education factor also explains roughly 3 percentage points although this contribution is less precisely estimated [27]

These results confirm the role that skills play as reported in the previous literature. Additionally, the finding that the parental valuation factor explains roughly 22 per cent of the gender gap points toward important impacts that parents can make during the teenage years. These parental impacts are much larger than what one might conclude from considering parental aspirations alone. In the simple Oaxaca-Blinder decompositions in Frenette and Zeman (2007), the parental aspirations variable, on its own, accounts for less than 10 per cent of the gap measured in the third cycle of YITS. However, parents also influence university attendance through their children's homework effort and grades. The factor model makes it possible to quantify, and aggregate the different channels through which parents' valuations operate.

## 5.3 Investigating why parental valuations are higher for girls?

The interpretation of the PVE factor as parents' valuations of education that are orthogonal to ability hinges on the restrictions imposed in the model. Key among those is the assumption that PISA reading scores are not a function of parental valuations. If there is a dimension of cognitive skill which is orthogonal to PISA reading skills but correlated with the parental measurement equations, then the PVE factor might simply reflect that unobserved ability. That would further imply that girls have higher PVE levels simply because they are more skilled in the unobserved dimension.

In the YITS data, there are two other test scores which I use to investigate this possibility. A

[27] Although, because of the non-linearities, the shares do not add up exactly to the whole explained portion, the discrepancy is quite small in practice.

random subset of the students, roughly half, wrote the PISA science test, while another subset wrote the PISA math test. If, after controlling for the reading scores, the PVE factor is correlated with the math or science scores, this would imply there is an important omitted skill biasing the PVE factor. I extract an estimated PVE factor for each sample member using Bayes Rule[28]:

$$\hat{\Theta} = \int \frac{p(Y|\hat{\Theta}, X, Z; \hat{\Gamma})p(\hat{\Theta}|X, Z; \hat{\Gamma})}{p(Y|X, Z)} d\hat{\Theta} \tag{24}$$

where $Y$ is a matrix of the enrollment outcome and all of the measurements, $\hat{\Gamma}$ is all of the estimated parameters in the model, and $\hat{\Theta}$ is the vector of the three estimated factors.

In Table 9, I report the results from a regression of Math and Science scores on the estimated PVE factor and the reading test scores. Unsurprisingly, the reading scores are highly correlated with both math and science scores. However, after controlling for those scores, the remaining variation in Science and Math scores is not statistically significantly related to the parental valuation factor. Indeed, the estimated coefficient is negative in all but the regression for girls' science scores. This evidence supports the claim that the PVE factor is not just another measure of skills.

If unobserved ability is not driving the results, why, then, are girls' parents more likely to highly value their daughters' education? In section 2.4, I suggest some potential reasons for gender differences in the PVE factor and now I look for suggestive evidence using responses to a question in the parental survey. After answering the aspirations question, parents were asked a follow-up question: "What is the main reason you hope your child will get this level of education?"[29] In Table 10, I present the distribution of answers given, separately for boys and girls and whether the parent had indicated 'university' or 'less than university' aspirations.

Although not directly informative about the unobserved PVE factor, this information does

[28] We also performed a similar exercise in FGG.

[29] In addition to the answers listed in the table, 'best choice in terms of financial costs' was also listed as an option. I included this choice in the 'other' category because it was chosen by very few parents. In the original full sample of 26,063, only 223 parents indicated costs were the main reason.

shed some light on what the parents had in mind while answering the aspirations question. Because the PVE factor loading is normalized to one in the parental aspirations question, 'university' aspirations means parents have a higher PVE. As such, when a reason is relatively more common among those with university aspirations, that reason will be correlated with higher parental valuations, and the converse is also true.

By far the most common reason given is 'better job opportunities or pay'. This reason is much more common among parents whose aspirations were less than university, which means it is correlated with lower parental valuations. This reason is also more common among the parents of boys. This pattern of responses is consistent with the idea that parents may value boys' university education less because their outside options have high rewards.

Turning to reasons that are positively correlated with parental valuations, only two reasons are more prevalent among the parents' with university aspirations. The first of these is 'Best match with child's ability', however, this reason is marginally more common among boys' parents. 'Valuable for personal growth and learning', is the single reason which is both more common among girls' parents and correlated with a higher valuation of education. Among those with university aspirations for their children, the parents of girls were 2.55 percentage points more likely to give this reason. While it is impossible to know what parents took this phrase to mean precisely, it certainly points toward the non-pecuniary benefits of education.

# 6 Conclusion

In the Youth in Transition Survey, 52% of girls and 37% of boys had ever attended university by age 21. I have sought to identify and quantify the underlying factors that contribute to that gender gap in university enrollment among Canadian youth. Using the factor model from Foley, Gallipoli, and Green (2014), I focus on three factors linked to cognitive skills, non-cognitive skills and parental valuations of education (PVE). I find that all three factors have large impacts on enrollment. The cognitive skill gradient is very large, raising the probability of attending university by roughly .5. The impact of non-cognitive skills and parental valuations can be almost as large but the impact depends on the level of the other factors. The impact of non-cognitive factors is larger when the parental valuation factor is lower. Similarly, the impact of the PVE factor is larger among less skilled youth.

All three factors also play an important role in explaining the gender gap in university attendance. At the highest cognitive skill level, the probability of attending university is virtually identical for boys and girls. Although gaps do persist among less skilled youth, overall the factor structure can account for 64 per cent of the total 15 percentage point gap. This is primarily because girls have higher levels of all three factors. The cognitive skill factor explains the largest fraction, but non-cognitive skills and parental valuations also play important roles. Indeed, the PVE factor explains 22.5 percent of the gap.

The results in this paper suggest that parents play a much larger role than one might expect from simply considering parents' education or their stated aspirations for their children. This is partly because the factor structure accounts for the extent to which variables, such as grades, are influenced not only by youth's skills and motivations, but also by how much their parents value education.

While contributing to a fuller understanding of why girls and boys differ in their propensity to attend university, this evidence also generates information that can be used when thinking about programs to promote university attendance, particularly among young men. As Fortin et al. (2015) point out, interventions such as 'Future to Discover' that offer information and financial assistance in the early years of high school show the potential to shift boys' plans for the future. My results also point toward models like the 'Future to Discover' intervention because it involved parents in the information component (Ford and Kwakye, 2016).

Involving parents is potentially important for two reasons. First, there is evidence that boys are less likely to make use of services (Angrist et al., 2009). Parental involvement might encourage boys to take advantage of available resources. The second reason goes beyond gender differences, to highlight the overall impact of parental valuations. To the extent that the parental valuation of education can be interpreted as something that is separable from ability and skills, these results suggest that parents' play a contemporaneous role in the university participation decision.

This was a key point that was emphasized in FGG in the context of the high school dropout decision. That point is worth restating here because of a critical difference between dropping out of high school and attending university. Among those with high cognitive skills, virtually everybody finishes high school in the YITS data. As such, in that group, parental valuations have no impact on dropping out. The same is not true for university participation. Even among the group of students with the highest cognitive and non-cognitive skills, parental valuations still have a large impact on the probability of attending university. Insofar as it is socially and economically desirable to encourage university participation among those with the highest level of skills, this is an issue that should be of relevance to policy makers.

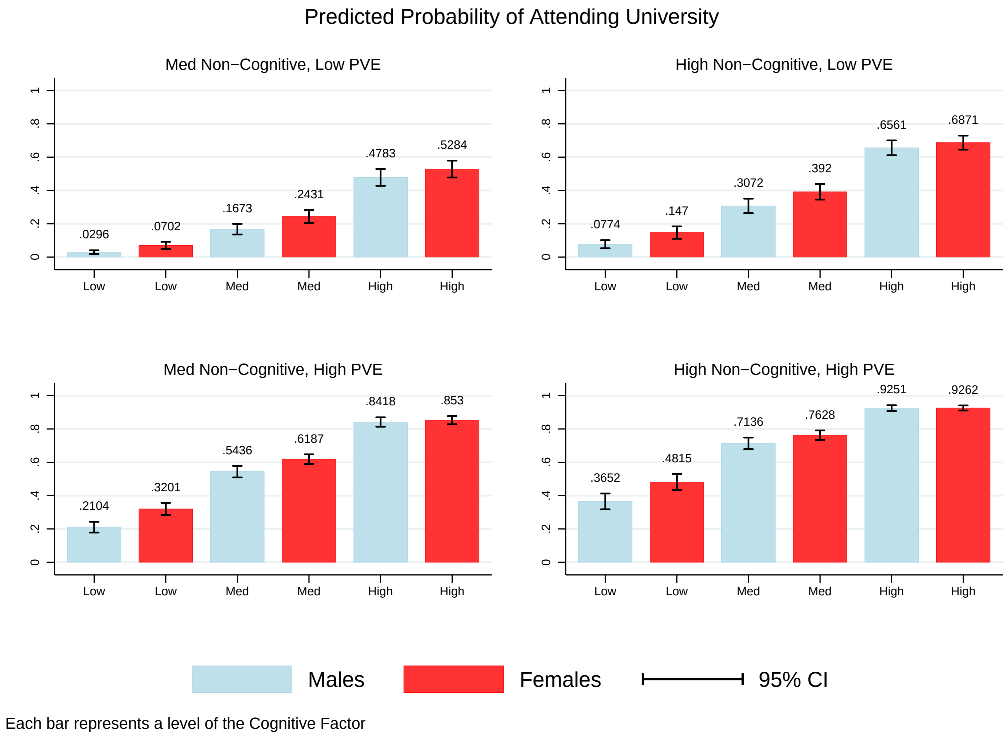


Figure 1: Predicted Probability of Attending University evaluated at each level of the estimated factors.

Notes: Confidence intervals constructed with standard errors estimated using the Delta Method and numerical derivatives.

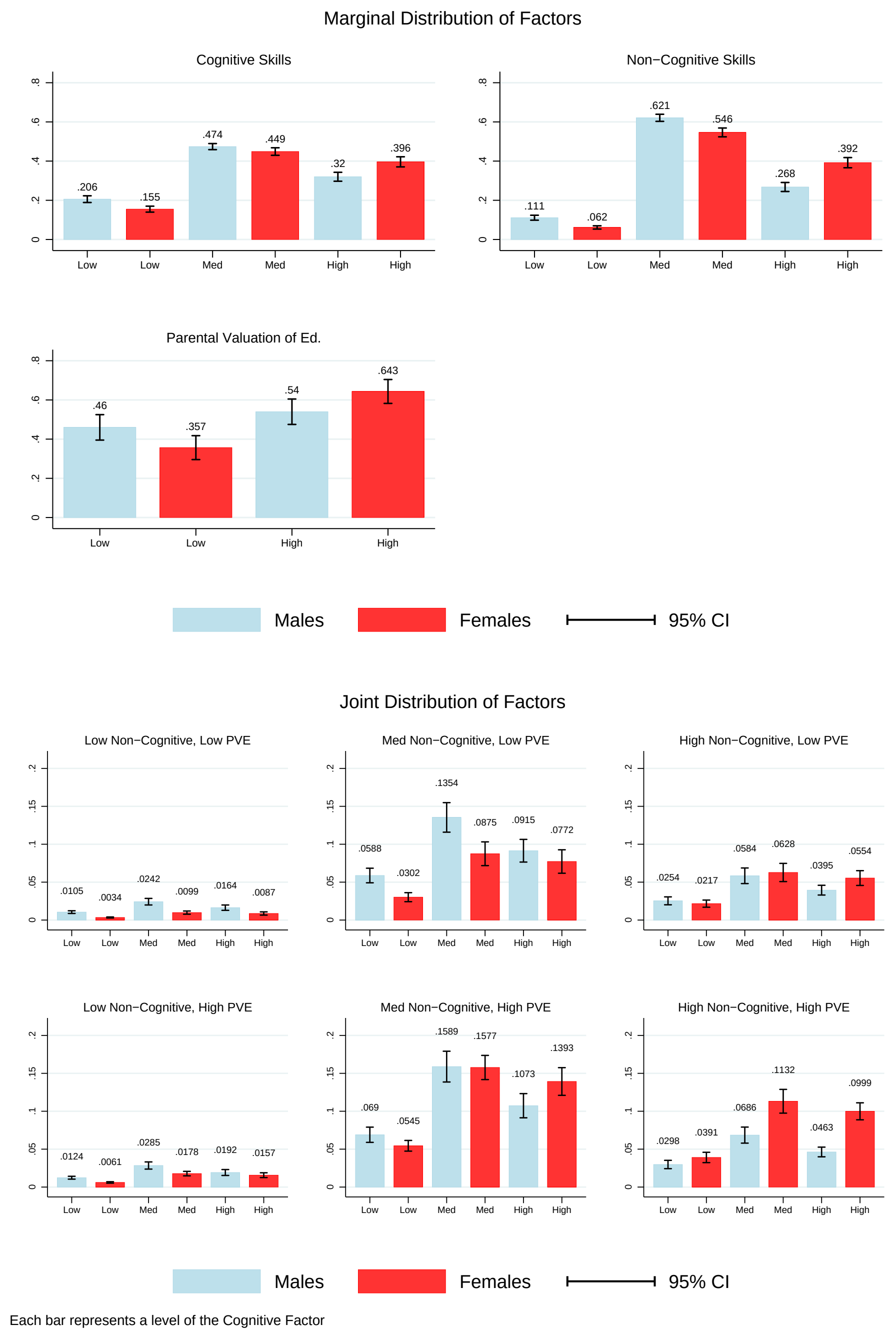


Figure 2: Estimated Factor Distributions

Notes: Confidence intervals constructed with standard errors estimated using the Delta Method and numerical derivatives.

Table 1: Outcome, Measurement, and Socioeconomics Variable Means (Standard deviations for continuous variables in parentheses)

| | **Pooled** | **Males** | **Females** |
|---|---|---|---|
| **Enrollment outcome and measurement variables** | | | |
| University enrollment | 0.4493 | 0.3747 | 0.5220 |
| Child wants university degree | 0.6430 | 0.5890 | 0.6957 |
| Parent hopes child gets degree | 0.6603 | 0.6182 | 0.7013 |
| **PISA quartiles** | | | |
| Quartile 1 (bottom) | 0.1938 | 0.2475 | 0.1415 |
| Quartile 2 | 0.2303 | 0.2422 | 0.2187 |
| Quartile 3 | 0.2767 | 0.2680 | 0.2851 |
| Quartile 4 (top) | 0.2992 | 0.2422 | 0.3547 |
| Overall grades (Percent) | 76.56 | 74.83 | 78.24 |
| | (10.65) | (10.71) | (10.31) |
| Child never just wants to get by | 0.3153 | 0.2266 | 0.4018 |
| Always does homework on time | 0.2566 | 0.2015 | 0.3103 |
| Parent saved for child's education | 0.6059 | 0.6051 | 0.6067 |
| **Gender and socioeconomic background** | | | |
| Female | 0.5065 | 0.0000 | 0.0000 |
| Adult Equivalent Family Income (1000) | 35,346 | 36,168 | 34,546 |
| | (26,875) | (28,485) | (25,183) |
| Ln of family income | 10.300 | 10.328 | 10.274 |
| | (0.606) | (0.597) | (0.613) |
| **Highest level of parental education** | | | |
| Both parents have less than HS | 0.0630 | 0.0625 | 0.0635 |
| One parent has HS | 0.1427 | 0.1410 | 0.1443 |
| Both parents have HS | 0.2031 | 0.2049 | 0.2013 |
| One parent has PSE below BA | 0.2329 | 0.2364 | 0.2294 |
| Both parents have PSE below BA | 0.1524 | 0.1551 | 0.1498 |
| One parent has BA or more | 0.1446 | 0.1429 | 0.1463 |
| Both parents have BA or more | 0.0614 | 0.0572 | 0.0655 |
| Indigenous | 0.0270 | 0.0284 | 0.0256 |
| Immigrant | 0.0785 | 0.0740 | 0.0828 |
| Lives in two parent family | 0.7306 | 0.7404 | 0.7210 |
| Rural | 0.2483 | 0.2445 | 0.2520 |
| **Province** | | | |
| Newfoundland | 0.0209 | 0.0195 | 0.0222 |
| Prince Edward Island | 0.0061 | 0.0057 | 0.0064 |
| Nova Scotia | 0.0350 | 0.0344 | 0.0355 |
| New Brunswick | 0.0272 | 0.0257 | 0.0288 |
| Quebec | 0.2285 | 0.2350 | 0.2221 |
| Ontario | 0.0369 | 0.0381 | 0.0357 |
| Manitoba | 0.0388 | 0.0389 | 0.0387 |
| Saskatchewan | 0.1078 | 0.1101 | 0.1055 |
| Alberta | 0.1260 | 0.1290 | 0.1230 |
| British Columbia | 0.3730 | 0.3637 | 0.3821 |
| Sample Size | 14,179 | 6,805 | 7,374 |

Table 2: Observed characteristics and university enrollment: Marginal Effects from Probit Regressions (Standard Errors in Parenthesis)

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Female | 0.150*** | 0.063*** | 0.115*** | 0.104*** | 0.049*** | 0.046*** |
| | (0.010) | (0.009) | (0.010) | (0.009) | (0.009) | (0.009) |
| **Highest level of parental education–Reference group Both less than HS** | | | | | | |
| One parent has HS | 0.062 | 0.048 | 0.059 | 0.055 | 0.045 | 0.046 |
| | (0.035) | (0.029) | (0.033) | (0.032) | (0.028) | (0.028) |
| Both parents have HS | 0.175*** | 0.103*** | 0.172*** | 0.112*** | 0.083*** | 0.082*** |
| | (0.029) | (0.025) | (0.028) | (0.027) | (0.024) | (0.024) |
| One parent has PSE below BA | 0.168*** | 0.094*** | 0.166*** | 0.105*** | 0.073** | 0.072** |
| | (0.028) | (0.025) | (0.028) | (0.026) | (0.024) | (0.024) |
| Both parents have PSE below BA | 0.256*** | 0.139*** | 0.256*** | 0.159*** | 0.108*** | 0.105*** |
| | (0.027) | (0.024) | (0.026) | (0.025) | (0.023) | (0.023) |
| One parent has BA or more | 0.385*** | 0.216*** | 0.374*** | 0.239*** | 0.168*** | 0.165*** |
| | (0.029) | (0.026) | (0.028) | (0.028) | (0.026) | (0.026) |
| Both parents have BA or more | 0.508*** | 0.287*** | 0.495*** | 0.319*** | 0.222*** | 0.219*** |
| | (0.029) | (0.027) | (0.028) | (0.027) | (0.026) | (0.026) |
| Lives in two parent family | 0.107*** | 0.054*** | 0.091*** | 0.084*** | 0.049*** | 0.047*** |
| | (0.013) | (0.012) | (0.013) | (0.012) | (0.011) | (0.011) |
| Log of family income | 0.059*** | 0.032** | 0.059*** | 0.022* | 0.014 | 0.015 |
| | (0.011) | (0.010) | (0.011) | (0.010) | (0.009) | (0.009) |
| Rural | -0.082*** | -0.079*** | -0.081*** | -0.036** | -0.053*** | -0.053*** |
| | (0.014) | (0.013) | (0.013) | (0.014) | (0.012) | (0.012) |
| Indigenous | -0.107** | -0.049 | -0.080* | -0.080** | -0.036 | -0.035 |
| | (0.033) | (0.030) | (0.032) | (0.030) | (0.028) | (0.028) |
| Immigrant | 0.070* | 0.069** | 0.043 | 0.005 | 0.024 | 0.022 |
| | (0.028) | (0.024) | (0.028) | (0.024) | (0.022) | (0.022) |
| **PISA Reading Test Scores–Reference group bottom quartile** | | | | | | |
| Q2 PISA Score | | 0.128*** | | | 0.094*** | 0.095*** |
| | | (0.017) | | | (0.016) | (0.016) |
| Q3 PISA Score | | 0.207*** | | | 0.156*** | 0.158*** |
| | | (0.017) | | | (0.016) | (0.016) |
| Q4 PISA Score | | 0.280*** | | | 0.220*** | 0.221*** |
| | | (0.017) | | | (0.017) | (0.017) |
| **Other measurements** | | | | | | |
| Child never just wants to get by | | | 0.098*** | | 0.016 | 0.014 |
| | | | (0.012) | | (0.010) | (0.010) |
| Always does homework on time | | | 0.163*** | | 0.052*** | 0.051*** |
| | | | (0.011) | | (0.010) | (0.010) |
| Child wants university degree | | | | 0.255*** | 0.142*** | 0.139*** |
| | | | | (0.011) | (0.011) | (0.011) |
| Parent hopes child gets degree | | | | 0.181*** | 0.104*** | 0.103*** |
| | | | | (0.011) | (0.010) | (0.010) |
| Parent saved for child's education | | | | 0.029** | 0.026** | 0.026** |
| | | | | (0.010) | (0.009) | (0.009) |
| Overall grades | | 0.014*** | | | 0.010*** | 0.010*** |
| | | (0.001) | | | (0.001) | (0.001) |
| **Variables not included in factor model** | | | | | | |
| Index of school quality | | | | | | -0.001 |
| | | | | | | (0.005) |
| Ratio of students to teachers in school | | | | | | -0.002 |
| | | | | | | (0.001) |
| Ratio of boys to girls in school | | | | | | 0.097 |
| | | | | | | (0.064) |
| Missing school information | | | | | | -0.006 |
| | | | | | | (0.038) |
| All friends think school important | | | | | | 0.026** |
| | | | | | | (0.010) |
| Sample Size | 14179 | 14179 | 14179 | 14179 | 14179 | 14179 |

Table 3: Coefficients from Factor Models, University Enrollment (Standard Errors in Parenthesis)

| | Restricted Model | Flexible Model | | |
|---|---|---|---|---|
| | | Common | Male | Female |
| Female | 0.3750*** | 0.2598 | | |
| | (0.0290) | (0.1639) | | |
| Lives in two parent family | 0.2384*** | 0.2374*** | | |
| | (0.0387) | (0.0384) | | |
| Rural | -0.3576*** | -0.3540*** | | |
| | (0.0402) | (0.0401) | | |
| Immigrant | 0.3325*** | 0.3326*** | | |
| | (0.0849) | (0.0849) | | |
| bmind | | | | |
| Ln of family income | 0.2362*** | 0.2373*** | | |
| | (0.0314) | (0.0307) | | |
| **Highest level of parental education–Reference group Both less than HS** | | | | |
| One parent has HS | 0.8750*** | 0.8552*** | | |
| | (0.0831) | (0.0826) | | |
| Both parents have HS | 1.1653*** | 1.1438*** | | |
| | (0.0777) | (0.0769) | | |
| One parent has PSE below BA | 1.1421*** | 1.1206*** | | |
| | (0.0727) | (0.0721) | | |
| Both parents have PSE below BA | 1.4363*** | 1.4206*** | | |
| | (0.0751) | (0.0742) | | |
| One parent has BA or more | 1.7682*** | 1.7481*** | | |
| | (0.0828) | (0.0821) | | |
| Both parents have BA or more | 2.1849*** | 2.1647*** | | |
| | (0.0887) | (0.0883) | | |
| Intercept | -5.1404*** | -5.0537*** | | |
| | (0.3320) | (0.3321) | | |
| Cognitive factor load ($\lambda^g_{0\theta_1}$) | 0.0170*** | | 0.0184*** | 0.0157*** |
| | (0.0006) | | (0.0008) | (0.0008) |
| Non-cognitive factor load ($\lambda^g_{0\theta_2}$) | 0.4845*** | | 0.4997*** | 0.4625*** |
| | (0.0346) | | (0.0423) | (0.0403) |
| PVE factor load ($\lambda^g_{0v_p}$) | 1.0281*** | | 1.0480*** | 0.9861*** |
| | (0.0697) | | (0.0841) | (0.0790) |

Table 4: Selected Coefficients from Factor Models, Measurement Equations (Standard Errors in Parenthesis)

| | **Restricted Model** | **Flexible Model** | | |
|---|---|---|---|---|
| | | **Common** | **Male** | **Female** |
| **yasp equation** | | | | |
| Intercept | -4.1039*** | -4.0279*** | | |
| | (0.3373) | (0.3485) | | |
| Cognitive factor load ($\lambda^g_{1\theta_1}$) | 0.0133*** | | 0.0141*** | 0.0129*** |
| | (0.0006) | | (0.0007) | (0.0009) |
| Non-cognitive factor load ($\lambda^g_{1\theta_2}$) | 0.4053*** | | 0.4164*** | 0.3863*** |
| | (0.0346) | | (0.0412) | (0.0435) |
| PVE factor load ($\lambda^g_{1v_p}$) | 1.2916*** | | 1.2583*** | 1.3590*** |
| | (0.0831) | | (0.0898) | (0.1093) |
| **PISA equation** | | | | |
| Intercept | 453.5429*** | 452.4679*** | | |
| | (13.6771) | (13.6918) | | |
| **getby equation** | | | | |
| Intercept | -2.2573*** | -2.2021*** | | |
| | (0.2490) | (0.2603) | | |
| PVE factor load ($\lambda^g_{3v_p}$) | 0.4419*** | | 0.4293*** | 0.4571*** |
| | (0.0518) | | (0.0657) | (0.0631) |
| **parasp equation** | | | | |
| Intercept | -3.6752*** | -3.6311*** | | |
| | (0.3079) | (0.3146) | | |
| Cognitive factor load ($\lambda^g_{4\theta_1}$) | 0.0103*** | | 0.0115*** | 0.0091*** |
| | (0.0005) | | (0.0006) | (0.0006) |
| Non-cognitive factor load ($\lambda^g_{4\theta_2}$) | 0.2587*** | | 0.2924*** | 0.2159*** |
| | (0.0264) | | (0.0334) | (0.0321) |
| Sample Size | 14179 | 14179 | 14179 | 14179 |

Table 5: Selected Coefficients from Factor Models, Measurement Equations (Standard Errors in Parenthesis) continued

| | Restricted Model | Flexible Model | | |
|---|---|---|---|---|
| | | Common | Male | Female |
| **grades equation** | | | | |
| Intercept | 56.2467*** | 56.2051*** | | |
| | (0.1431) | (0.1516) | | |
| Cognitive factor load ($\lambda^g_{5\theta_1}$) | 0.1708*** | | 0.1708*** | 0.1713*** |
| | (0.0031) | | (0.0032) | (0.0032) |
| Non-cognitive factor load ($\lambda^g_{5\theta_2}$) | 9.3205*** | | 9.3037*** | 9.2621*** |
| | (0.5437) | | (0.5445) | (0.5426) |
| PVE factor load ($\lambda^g_{5v_p}$) | 0.0554*** | | 0.0590*** | 0.0501*** |
| | (0.0140) | | (0.0220) | (0.0163) |
| **hmwrk equation** | | | | |
| Intercept | | | | |
| Non-cognitive factor load ($\lambda^g_{6\theta_2}$) | 0.6091*** | | 0.6055*** | 0.6117*** |
| | (0.0368) | | (0.0474) | (0.0438) |
| PVE factor load ($\lambda^g_{6v_p}$) | 0.4558*** | | 0.3869*** | 0.5136*** |
| | (0.0590) | | (0.0781) | (0.0713) |
| **saved equation** | | | | |
| Intercept | -4.9903*** | -4.8697*** | | |
| | (0.3129) | (0.3162) | | |
| Cognitive factor load ($\lambda^g_{7\theta_1}$) | 0.0007** | | 0.0011** | 0.0003 |
| | (0.0003) | | (0.0005) | (0.0005) |
| Non-cognitive factor load ($\lambda^g_{7\theta_2}$) | 0.0758*** | | 0.0544*** | 0.1030*** |
| | (0.0164) | | (0.0210) | (0.0247) |
| PVE factor load ($\lambda^g_{5v_p}$) | 0.2440*** | | 0.1755*** | 0.3104*** |
| | (0.0300) | | (0.0418) | (0.0416) |
| **Factor' locations** | | | | |
| Low cognitive factor location | -117.2557*** | -117.0948*** | | |
| | (2.1390) | (2.1572) | | |
| Medium cognitive factor location | -58.4432*** | -58.3647*** | | |
| | (1.0723) | (1.0807) | | |
| High non-cognitive factor location | 2.0589*** | 2.0653*** | | |
| | (0.1200) | (0.1207) | | |
| Medium non-cognitive factor location | 3.1392*** | 3.1503*** | | |
| | (0.1829) | (0.1841) | | |
| High PVE factor location | 1.2087*** | 1.2088*** | | |
| | (0.0566) | (0.0571) | | |
| Sample Size | 14179 | 14179 | 14179 | 14179 |
| Log likelihood | -99459.209 | -99237.838 | 0.000 | 0.000 |

Table 5: Marginal Effect of Cognitive Factor

| | **Low Non-Cognitive** | | **Medium Non-Cognitive** | | **High Non-Cognitive** | |
|---|---|---|---|---|---|---|
| | Low PVE | High PVE | Low PVE | High PVE | Low PVE | High PVE |
| **Marginal Effect of High to Low** | | | | | | |
| Males | -0.1734*** | -0.5112*** | -0.4487*** | -0.6314*** | -0.5786*** | -0.5599*** |
| | (0.0251) | (0.0305) | (0.0236) | (0.0205) | (0.0204) | (0.0232) |
| Females | -0.2214*** | -0.5044*** | -0.4582*** | -0.5329*** | -0.5401*** | -0.4446*** |
| | (0.0282) | (0.0271) | (0.0226) | (0.0216) | (0.0205) | (0.0236) |
| **Marginal Effect of Medium to Low** | | | | | | |
| Males | -0.1437*** | -0.3349*** | -0.3110*** | -0.2982*** | -0.3488*** | -0.2115*** |
| | (0.0194) | (0.0176) | (0.0157) | (0.0110) | (0.0141) | (0.0117) |
| Females | -0.1662*** | -0.2961*** | -0.2853*** | -0.2343*** | -0.2951*** | -0.1634*** |
| | (0.0196) | (0.0146) | (0.0144) | (0.0097) | (0.0138) | (0.0096) |

Table 6: Marginal Effect of Non-Cognitive Factor

| | **Low Cognitive** | | **Medium Cognitive** | | **High Cognitive** | |
|---|---|---|---|---|---|---|
| | Low PVE | High PVE | Low PVE | High PVE | Low PVE | High PVE |
| **Marginal Effect of High to Low** | | | | | | |
| Males | -0.0747*** | -0.3193*** | -0.2748*** | -0.4914*** | -0.4799*** | -0.3680*** |
| | (0.0118) | (0.0233) | (0.0198) | (0.0273) | (0.0247) | (0.0318) |
| Females | -0.1359*** | -0.3801*** | -0.3257*** | -0.4531*** | -0.4546*** | -0.3204*** |
| | (0.0173) | (0.0251) | (0.0210) | (0.0283) | (0.0262) | (0.0301) |
| **Marginal Effect of Medium to Low** | | | | | | |
| Males | -0.0478*** | -0.1548*** | -0.1399*** | -0.1700*** | -0.1778*** | -0.0833*** |
| | (0.0072) | (0.0124) | (0.0108) | (0.0104) | (0.0117) | (0.0076) |
| Females | -0.0768*** | -0.1614*** | -0.1489*** | -0.1441*** | -0.1587*** | -0.0731*** |
| | (0.0095) | (0.0119) | (0.0109) | (0.0093) | (0.0115) | (0.0067) |

Table 7: Marginal Effect of PVE factors

| | **Low Cognitive** | | | **Medium Cognitive** | | | **High Cognitive** | | |
|---|---|---|---|---|---|---|---|---|---|
| Non-Cog: | Low | Med | High | Low | Med | High | Low | Med | High |
| **Marginal Effect of High to Low** | | | | | | | | | |
| Males | -0.0432*** | -0.1898*** | -0.3810*** | -0.1808*** | -0.3763*** | -0.3635*** | -0.2878*** | -0.4064*** | -0.2691*** |
| | (0.0071) | (0.0174) | (0.0217) | (0.0134) | (0.0192) | (0.0224) | (0.0186) | (0.0221) | (0.0197) |
| Females | -0.0903*** | -0.2433*** | -0.3733*** | -0.2499*** | -0.3756*** | -0.3246*** | -0.3345*** | -0.3708*** | -0.2391*** |
| | (0.0106) | (0.0173) | (0.0211) | (0.0145) | (0.0208) | (0.0228) | (0.0190) | (0.0230) | (0.0192) |

Table 8: Decomposition of Gender Gap in University Enrollment

| | | Male Parameters | | | Female Parameters | | |
|---|---|---|---|---|---|---|---|
| | | **Explained by** | | **Unexplained** | **Explained by** | | **Unexplained** |
| | | Observed Var | Parameter(s) | | Observed Var | Parameter(s) | |
| Row | **Predicted gap: .1508** | | | | | | |
| (1) | Observed Characteristics | -0.0079 | | 0.1587 | -0.0093 | | 0.1601 |
| | | (0.0031) | | (0.0080) | (0.0034) | | (0.0080) |
| (2) | Full factor structure | | 0.0960 | 0.0548 | | 0.0934 | 0.0574 |
| | | | (0.0391) | (0.0390) | | (0.0423) | (0.0421) |
| **Factor loading** | | | | | | | |
| (3) | All 3 | | -0.0011 | 0.0972 | | 0.0011 | 0.0923 |
| | | | (0.0351) | (0.0151) | | (0.0348) | (0.0156) |
| (4) | Cognitive | | 0.0315 | 0.0645 | | 0.0324 | 0.0609 |
| | | | (0.0121) | (0.0382) | | (0.0130) | (0.0412) |
| (5) | Non-cognitive | | -0.0217 | 0.1177 | | -0.0196 | 0.1130 |
| | | | (0.0278) | (0.0234) | | (0.0252) | (0.0262) |
| (6) | VPE | | -0.0115 | 0.1075 | | -0.0105 | 0.1039 |
| | | | (0.0155) | (0.0336) | | (0.0143) | (0.0362) |
| **Factor distribution** | | | | | | | |
| (7) | All 3 | | 0.0921 | 0.0039 | | 0.0974 | -0.0040 |
| | | | (0.0153) | (0.0326) | | (0.0153) | (0.0383) |
| (8) | Cognitive | | 0.0308 | 0.0652 | | 0.0343 | 0.0590 |
| | | | (0.0055) | (0.0390) | | (0.0064) | (0.0414) |
| (9) | Non-cognitive | | 0.0280 | 0.0680 | | 0.0281 | 0.0652 |
| | | | (0.0034) | (0.0385) | | (0.0038) | (0.0436) |
| (10) | VPE | | 0.0334 | 0.0627 | | 0.0332 | 0.0602 |
| | | | (0.0148) | (0.0335) | | (0.0142) | (0.0379) |

Table 9: Correlation between PVE and PISA Math and Science Scores

| | Math | | Science | |
|---|---|---|---|---|
| | **Males** | **Females** | **Males** | **Females** |
| PISA reading scores | 0.7410*** | 0.7329*** | 0.8653*** | 0.8520*** |
| | (0.0106) | (0.0117) | (0.0087) | (0.0099) |
| Parental Valuation Factor | -0.1967 | -0.4433 | -0.9762 | 0.0223 |
| | (0.7991) | (0.7883) | (0.6191) | (0.7600) |
| Sample Size | 3737 | 4151 | 3720 | 4053 |

Table 10: Main reasons for Parental Aspirations

| | **< University Aspirations** | | **University Aspirations** | |
|---|---|---|---|---|
| | **Boys** | **Girls** | **Boys** | **Girls** |
| Better job opportunities or pay | 0.6006 | 0.5635 | 0.5363 | 0.4768 |
| | (0.0095) | (0.0106) | (0.0075) | (0.0069) |
| Valuable for personal growth and learning | 0.1027 | 0.1231 | 0.1562 | 0.1741 |
| | (0.0068) | (0.0076) | (0.0054) | (0.0050) |
| Child's choice | 0.1028 | 0.1190 | 0.0563 | 0.0903 |
| | (0.0054) | (0.0060) | (0.0042) | (0.0039) |
| Best match with child's ability | 0.0840 | 0.0751 | 0.1194 | 0.1136 |
| | (0.0059) | (0.0065) | (0.0046) | (0.0043) |
| Other | 0.1099 | 0.1194 | 0.1318 | 0.1452 |
| | (0.0065) | (0.0072) | (0.0051) | (0.0047) |

# References


M. Almlund, A. L. Duckworth, J. Heckman, and T. Kautz. *Personality Psychology and Economics*, volume 4 of *Handbook of the Economics of Education*, chapter 0, pages 1–181. Elsevier, May/June 2011.

J. Angrist, D. Lang, and P. Oreopoulos. Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, 1(1):136–163, 2009.

D. Autor, D. Figlio, K. Karbownik, J. Roth, and M. Wasserman. Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes. NBER Working Papers 22267, National Bureau of Economic Research, Inc, May 2016.

M. Baker and K. Milligan. Boy-Girl Differences in Parental Time Investments: Evidence from Three Countries. *Journal of Human Capital*, 10(4):399–441, 2016.

G. S. Becker. A Theory of Marriage: Part I. *Journal of Political Economy*, 81(4):813–46, July-Aug. 1973.

G. S. Becker, W. H. J. Hubbard, and K. M. Murphy. Explaining the worldwide boom in higher education of women. *Journal of Human Capital*, 4(3):203–241, 2010.

M. Bertrand and J. Pan. The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior. *American Economic Journal: Applied Economics*, 5(1):32–64, January 2013.

A. Bisin and T. Verdier. The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory*, 97(2):298–319, April 2001.

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

L. Borghans, B. H. Golsteyn, J. J. Heckman, and J. E. Humphries. Identification Problems in Personality Psychology. NBER Working Papers 16917, National Bureau of Economic Research, Inc, Mar. 2011.

B. Boudarbat, T. Lemieux, and W. C. Riddell. The Evolution of the Returns to Human Capital in Canada, 1980-2005. *Canadian Public Policy*, 36(1):63–89, March 2010.

A. A. Brenøe and S. Lundberg. Gender gaps in the effects of childhood family environment: Do they persist into adulthood? *European Economic Review*, 2017. doi: https://doi.org/10.1016/j.euroecorev.2017.04.004.

C. Buchmann and T. A. DiPrete. The growing female advantage in college completion: The role of family background and academic achievement. *American Sociological Review*, 71(4): 515–541, 2006.

S. V. Cameron and C. Taber. Estimation of educational borrowing constraints using returns to schooling. *Journal of Political Economy*, 112(1):132–182, 2004.

V. Caponi and M. Plesca. Post-secondary education in Canada: can ability bias explain the earnings gap between college and university graduates? *Canadian Journal of Economics*, 42 (3):1100–1131, August 2009.

D. Card, A. Payne, and C. Sechel. Understanding the gender gap in university participation: An exploration of the application behaviour on ontario high school students. Technical report, Higher Education Quality Council of Ontario., Toronto, 2011.

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

R. Chetty, N. Hendren, F. Lin, J. Majerovitz, and B. Scuderi. Childhood Environment and Gender Gaps in Adulthood. *American Economic Review*, 106(5):282–288, May 2016.

P.-A. Chiappori, M. Iyigun, and Y. Weiss. Investment in Schooling and the Marriage Market. *American Economic Review*, 99(5):1689–1713, December 2009.

P.-A. Chiappori, B. Salanié, and Y. Weiss. Partner Choice and the Marital College Premium: Analyzing Marital Patterns Over Several Decades. *American Economic Review*, forthcoming.

L. N. Christofides, M. Hoy, and Z. Li. Evolution of aspirations for university attendance A gender comparison. A MESA Project Research Paper. MESA2008-3, Educational Policy Institute, Toronto, 2008a.

L. N. Christofides, M. Hoy, and Z. Li. Evolution of aspirations for university attendance. In R. Finnie, R. E. Meuller, A. Sweetman, and A. Usher, editors, *Who Goes? Who stays? What matters?: Accessing and Persisting in Post-Secondary Education in Canada*, pages 109–134. McGill-Queens University Press, 2008b.

L. N. Christofides, M. Hoy, and L. Yang. The determinants of university participation in canada (1977–2003). *Canadian Journal of Higher Education*, 39(2):1–24, 2009.

L. N. Christofides, M. Hoy, and L. Yang. Participation in Canadian Universities: The gender imbalance (1977-2005). *Economics of Education Review*, 29(3):400–410, June 2010.

D. Conger and M. C. Long. Why are men falling behind? gender gaps in college performance and persistence. *The Annals of the American Academy of Political and Social Science*, 627 (1):184–214, 2010.

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

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

F. Cunha, J. Heckman, and S. Navarro. Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers*, 57(2):191–261, April 2005.

F. Cunha, J. J. Heckman, and S. M. Schennach. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931, 05 2010.

F. Cunha, I. Elo, and J. Culhane. Eliciting Maternal Expectations about the Technology of Cognitive Skill Formation. NBER Working Papers 19144, National Bureau of Economic Research, Inc, June 2013.

T. S. Dee. A Teacher Like Me: Does Race, Ethnicity, or Gender Matter? *American Economic Review*, 95(2):158–165, May 2005.

M. Drolet. Participation in post-secondary education in canada: Has the role changed over the 1990s? *No 11F0019MIE Statistics Canada, Analytical Studies Branch*, (243), 2005.

C. Echevarria and A. Merlo. Gender Differences in Education in a Dynamic Household Bargaining Model. *International Economic Review*, 40(2):265–86, May 1999.

R. W. Fairlie. The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment. *Journal of Labor Economics*, 17(1):80–108, January 1999.

C. S. Finnie, R. and A. Wismer. Under-represented groups in postsecondary education in ontario: Evidence from the youth in transition survey. Technical report, Higher Education Quality Council of Ontario., Toronto, 2011.

K. Foley, G. Gallipoli, and D. A. Green. Ability, Parental Valuation of Education, and the High School Dropout Decision. *Journal of Human Resources*, 49(4):906–944, 2014.

R. Ford and I. Kwakye. Future to discover: Sixth year post-secondary impacts repor. Technical report, Social Research and Demonostration Corperation, July 2016.

N. Fortin, T. Lemieux, and S. Firpo. *Decomposition Methods in Economics*, volume 4 of *Handbook of Labor Economics*, chapter 1, pages 1–102. Elsevier, 2011.

N. M. Fortin, P. Oreopoulos, and S. Phipps. Leaving Boys Behind: Gender Disparities in High Academic Achievement. *Journal of Human Resources*, 50(3):549–579, 2015.

M. Frenette and K. Zeman. Why Are Most University Students Women? Evidence Based on Academic Performance, Study Habits and Parental Influences. Analytical Studies Branch Research Paper Series 2007303e, Statistics Canada, Analytical Studies Branch, Sept. 2007.

C. Goldin. The meaning of college in the lives of american women: The past one-hundred years. Working Paper 4099, National Bureau of Economic Research, June 1992.

C. Goldin. Career and Family: College Women Look to the Past. NBER Working Papers 5188, National Bureau of Economic Research, Inc, July 1995.

C. Goldin, L. F. Katz, and I. Kuziemko. The homecoming of american college women: The reversal of the college gender gap. *The Journal of Economic Perspectives*, 20(4):133–133, 2006.

Z. Griliches and W. M. Mason. Education, Income, and Ability. *Journal of Political Economy*, 80(3):74–103, 1972.

D. F. Halpern. *Sex Differences in Cognitive Abilities*. Psychology Press, New York, fourth edition, 2012.

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

J. J. Heckman and S. Mosso. The Economics of Human Development and Social Mobility. *Annual Review of Economics*, 6(1):689–733, 08 2014.

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

A. J. Hill. The Girl Next Door: The Effect of Opposite Gender Friends on High School Achievement. *American Economic Journal: Applied Economics*, 7(3):147–77, July 2015.

A. J. Houtenville and K. S. Conway. Parental Effort, School Resources, and Student Achievement. *Journal of Human Resources*, 43(2):437–453, 2008.

C. Hoxby. Peer Effects in the Classroom: Learning from Gender and Race Variation. NBER Working Papers 7867, National Bureau of Economic Research, Inc, Aug. 2000.

B. A. Jacob. Where the boys aren't: non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education Review*, 21(6):589–598, 2002.

S. Junor and A. Usher. *The price of knowledge 2004: access and student finance in Canada/written by Sean Junor, Alex Usher.* Montreal, Que.: Canada Millennium Scholarship Foundation,, 2004.

R. R. McCrae and O. P. John. An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2):175–215, 1992. ISSN 1467-6494.

A. Motte, H. Qiu, Y. Zhang, and P. Bussière. The youth in transition survey: Following canadian youth through time. In R. Finnie, R. E. Meuller, A. Sweetman, and A. Usher, editors, *Who Goes? Who stays? What matters?: Accessing and Persisting in Post-Secondary Education in Canada*, pages 63–75. McGill-Queens University Press, 2008.

OECD. Education at a glance 2015. Technical report, OECD Publishing, 2016.

P. Oreopoulos and K. G. Salvanes. Priceless: The Nonpecuniary Benefits of Schooling. *Journal of Economic Perspectives*, 25(1):159–84, Winter 2011.

M. Peters and A. Siow. Competing Pre-marital Investments. Working Papers peters-00-01, University of Toronto, Department of Economics, Jan. 2000.

M. Turcotte. Women and education. Women in Canada: a Gender-based Statistical Report, 6th Edition, Statistics Canada, 2012.

# A Predicted probabilities for all factor points of support

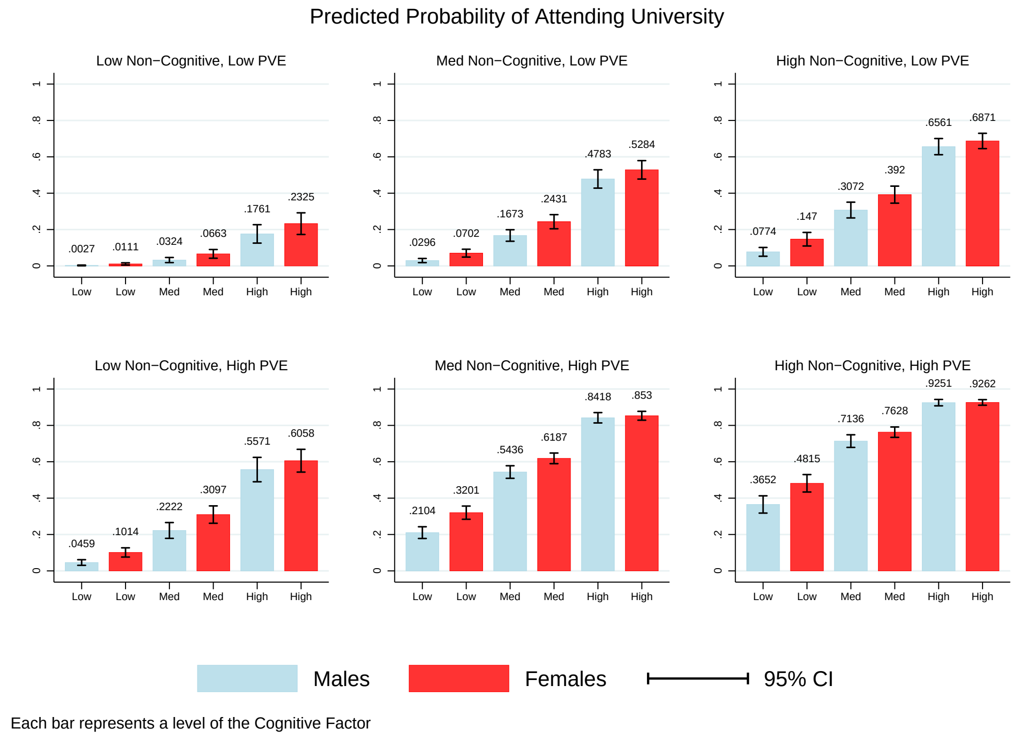


Figure 3: Predicted Probability of Attending University evaluated at each level of the estimated factors.

Notes: Confidence intervals constructed with standard errors estimated using the Delta Method and numerical derivatives.

# B Detailed Decompositions

This appendix describes the detailed decompositions which are reported in columns 2 through 11 in Table 8

The predicted probabilities for males and females are, respectively:

$$
\begin{aligned}
U\left(X^{m}, m\right) &= \frac{1}{n^{m}} \sum_{i=1}^{n^{m}} \sum_{\theta_1} \sum_{\theta_2} \sum_{v_p} p^{m}\left(\theta_1\right) p^{m}\left(\theta_2\right) p^{m}\left(v_p\right) F\left(\gamma_0^{m} + \gamma_x X_i^{m} + \lambda_{0\theta_1}^{m} \theta_{i1} + \lambda_{0\theta_2}^{m} \theta_{i2} + \lambda_{0v_p}^{m} v_{ip}\right) \\
U\left(X^{f}, f\right) &= \frac{1}{n^{f}} \sum_{i=1}^{n^{f}} \sum_{\theta_1} \sum_{\theta_2} \sum_{v_p} p^{f}\left(\theta_1\right) p^{f}\left(\theta_2\right) p^{f}\left(v_p\right) F\left(\gamma_0^{f} + \gamma_x X_i^{f} + \lambda_{0\theta_1}^{f} \theta_{i1} + \lambda_{0\theta_2}^{f} \theta_{i2} + \lambda_{0v_p}^{f} v_{ip}\right)
\end{aligned}
$$

The counterfactual for decomposing the fraction of the gap attributable to the entire factor structure is equation (21):

$$
U\left(X^{g}, \gamma_0^{g}, h\right) = \frac{1}{n^{g}} \sum_{i=1}^{n^{g}} \sum_{\theta_1} \sum_{\theta_2} \sum_{v_p} p^{h}\left(\theta_1\right) p^{h}\left(\theta_2\right) p^{h}\left(v_p\right) F\left(\gamma_0^{g} + \gamma_x X_i^{g} + \lambda_{0\theta_1}^{h} \theta_{i1} + \lambda_{0\theta_2}^{h} \theta_{i2} + \lambda_{0v_p}^{h} v_{ip}\right)
$$

**The decomposition in Row 2 is:**

Using the male parameters:

$$
\Delta_{\Lambda\Theta^{m}} = \underbrace{U\left(X^{f}, f\right) - U\left(X^{f}, \gamma_0^{f}, m\right)}_{\text{Explained by factor structure}} + \underbrace{U\left(X^{f}, \gamma_0^{f}, m\right) - U\left(X^{m}, m\right)}_{\text{Unexplained}}
$$

Using the female parameters:

$$
\Delta_{\Lambda\Theta^{f}} = \underbrace{U\left(X^{m}, \gamma_0^{m}, f\right) - U\left(X^{m}, m\right)}_{\text{Explained by factor structure}} + \underbrace{U\left(X^{f}, f\right) - U\left(X^{m}, \gamma_0^{m}, f\right)}_{\text{Unexplained}}
$$

**The counterfactual in Row 4 is:**

$$
U\left(g, \lambda_{0\theta_1}^{h}\right) = \frac{1}{n^{g}} \sum_{i=1}^{n^{g}} \sum_{\theta_1} \sum_{\theta_2} \sum_{v_p} p^{h}\left(\theta_1\right) p^{g}\left(\theta_2\right) p^{g}\left(v_p\right) F\left(\gamma_0^{g} + \gamma_x X_i^{g} + \lambda_{0\theta_1}^{h} \theta_{i1} + \lambda_{0\theta_2}^{g} \theta_{i2} + \lambda_{0v_p}^{g} v_{ip}\right)
$$

**The decomposition in Row 4 is:**

Using the male parameters:

$$
U\left(X^{f}, f\right) - U\left(X^{f}, \gamma_0^{f}, m\right) = \underbrace{U\left(X^{f}, f\right) - U\left(f, \lambda_{0\theta_1}^{m}\right)}_{\text{Explained by } \lambda_{0\theta_1}^{m}} + \underbrace{U\left(f, \lambda_{0\theta_1}^{m}\right) - U\left(X^{f}, \gamma_0^{f}, m\right)}_{\text{Unexplained}}
$$

Using the female parameters:

$$
U\left(X^{m}, \gamma_0^{m}, f\right) - U\left(X^{m}, m\right) = \underbrace{U\left(m, \lambda_{0\theta_1}^{f}\right) - U\left(X^{m}, m\right)}_{\text{Explained by } \lambda_{0\theta_1}^{f}} + \underbrace{U\left(X^{m}, \gamma_0^{m}, f\right) - U\left(m, \lambda_{0\theta_1}^{f}\right)}_{\text{Unexplained}}
$$

**The counterfactual in Row 5 is:**

$$U\left(g,\lambda^h_{0\theta_2}\right) = \frac{1}{n^g}\sum_{i=1}^{n^g}\sum_{\theta_1}\sum_{\theta_2}\sum_{v_p} p^g(\theta_1)\,p^g(\theta_2)\,p^g(v_p)\,F\left(\gamma^g_0+\gamma_x X^g_i+\lambda^g_{0\theta_1}\theta_{i1}+\lambda^h_{0\theta_2}\theta_{i2}+\lambda^g_{0v_p}v_{ip}\right)$$

**The decomposition in Row 5 is:**

Using the male parameters:

$$U\left(X^f,f\right)-U\left(X^f,\gamma^f_0,m\right) = \underbrace{U\left(X^f,f\right)-U\left(f,\lambda^m_{0\theta_2}\right)}_{\text{Explained by }\lambda^m_{0\theta_2}} + \underbrace{U\left(f,\lambda^m_{0\theta_2}\right)-U\left(X^f,\gamma^f_0,m\right)}_{\text{Unexplained}}$$

Using the female parameters:

$$U\left(X^m,\gamma^m_0,f\right)-U\left(X^m,m\right) = \underbrace{U\left(m,\lambda^f_{0\theta_2}\right)-U\left(X^m,m\right)}_{\text{Explained by }\lambda^f_{0\theta_2}} + \underbrace{U\left(X^m,\gamma^m_0,f\right)-U\left(m,\lambda^f_{0\theta_2}\right)}_{\text{Unexplained}}$$

**The counterfactual in Row 6 is:**

$$U\left(g,\lambda^h_{0v_p}\right) = \frac{1}{n^g}\sum_{i=1}^{n^g}\sum_{\theta_1}\sum_{\theta_2}\sum_{v_p} p^g(\theta_1)\,p^g(\theta_2)\,p^g(v_p)\,F\left(\gamma^g_0+\gamma_x X^g_i+\lambda^g_{0\theta_1}\theta_{i1}+\lambda^g_{0\theta_2}\theta_{i2}+\lambda^h_{0v_p}v_{ip}\right)$$

**The decomposition in Row 6 is:**

Using the male parameters:

$$U\left(X^f,f\right)-U\left(X^f,\gamma^f_0,m\right) = \underbrace{U\left(X^f,f\right)-U\left(f,\lambda^m_{0v_p}\right)}_{\text{Explained by }\lambda^m_{0v_p}} + \underbrace{U\left(f,\lambda^m_{0v_p}\right)-U\left(X^f,\gamma^f_0,m\right)}_{\text{Unexplained}}$$

Using the female parameters:

$$U\left(X^m,\gamma^m_0,f\right)-U\left(X^m,m\right) = \underbrace{U\left(m,\lambda^f_{0v_p}\right)-U\left(X^m,m\right)}_{\text{Explained by }\lambda^f_{0v_p}} + \underbrace{U\left(X^m,\gamma^m_0,f\right)-U\left(m,\lambda^f_{0v_p}\right)}_{\text{Unexplained}}$$

**The counterfactual in Row 8is:**

$$U\left(g,p^h(\theta_1)\right) = \frac{1}{n^g}\sum_{i=1}^{n^g}\sum_{\theta_1}\sum_{\theta_2}\sum_{v_p} p^h(\theta_1)\,p^g(\theta_2)\,p^g(v_p)\,F\left(\gamma^g_0+\gamma_x X^g_i+\lambda^g_{0\theta_1}\theta_{i1}+\lambda^g_{0\theta_2}\theta_{i2}+\lambda^g_{0v_p}v_{ip}\right)$$

**The decomposition in Row 8 is:**

Using the male parameters:

$$U\left(X^f,f\right)-U\left(X^f,\gamma^f_0,m\right) = \underbrace{U\left(X^f,f\right)-U\left(f,p^m(\theta_1)\right)}_{\text{Explained by }p^m(\theta_1)} + \underbrace{U\left(f,p^m(\theta_1)\right)-U\left(X^f,\gamma^f_0,m\right)}_{\text{Unexplained}}$$

Using the female parameters:

$$U\left(X^m,\gamma^m_0,f\right)-U\left(X^m,m\right) = \underbrace{U\left(m,p^f(\theta_1)\right)-U\left(X^m,m\right)}_{\text{Explained by }p^f(\theta_1)} + \underbrace{U\left(X^m,\gamma^m_0,f\right)-U\left(m,p^f(\theta_1)\right)}_{\text{Unexplained}}$$

**The counterfactual in Row 9 is:**

$$U\left(g,p^{h}\left(\theta_{2}\right)\right) = \frac{1}{n^{g}}\sum_{i=1}^{n^{g}}\sum_{\theta_{1}}\sum_{\theta_{2}}\sum_{v_{p}}p^{g}\left(\theta_{1}\right)p^{h}\left(\theta_{2}\right)p^{g}\left(v_{p}\right)F\left(\gamma_{0}^{g}+\gamma_{x}X_{i}^{g}+\lambda_{0\theta_{1}}^{g}\theta_{i1}+\lambda_{0\theta_{2}}^{g}\theta_{i2}+\lambda_{0v_{p}}^{g}v_{ip}\right)$$

**The decomposition in Row 9 is:**

Using the male parameters:

$$U\left(X^{f},f\right)-U\left(X^{f},\gamma_{0}^{f},m\right) = \underbrace{U\left(X^{f},f\right)-U\left(f,p^{m}\left(\theta_{2}\right)\right)}_{\text{Explained by } p^{m}\left(\theta_{2}\right)} + \underbrace{U\left(f,p^{m}\left(\theta_{2}\right)\right)-U\left(X^{f},\gamma_{0}^{f},m\right)}_{\text{Unexplained}}$$

Using the female parameters:

$$U\left(X^{m},\gamma_{0}^{m},f\right)-U\left(X^{m},m\right) = \underbrace{U\left(m,p^{f}\left(\theta_{2}\right)\right)-U\left(X^{m},m\right)}_{\text{Explained by } p^{f}\left(\theta_{2}\right)} + \underbrace{U\left(X^{m},\gamma_{0}^{m},f\right)-U\left(m,p^{f}\left(\theta_{2}\right)\right)}_{\text{Unexplained}}$$

**The counterfactual in Row 10 is:**

$$U\left(g,p^{h}\left(v_{p}\right)\right) = \frac{1}{n^{g}}\sum_{i=1}^{n^{g}}\sum_{\theta_{1}}\sum_{\theta_{2}}\sum_{v_{p}}p^{g}\left(\theta_{1}\right)p^{g}\left(\theta_{2}\right)p^{h}\left(v_{p}\right)F\left(\gamma_{0}^{g}+\gamma_{x}X_{i}^{g}+\lambda_{0\theta_{1}}^{g}\theta_{i1}+\lambda_{0\theta_{2}}^{g}\theta_{i2}+\lambda_{0v_{p}}^{g}v_{ip}\right)$$

**The decomposition in Row 10 is:**

Using the male parameters:

$$U\left(X^{f},f\right)-U\left(X^{f},\gamma_{0}^{f},m\right) = \underbrace{U\left(X^{f},f\right)-U\left(f,p^{m}\left(v_{p}\right)\right)}_{\text{Explained by } p^{m}\left(v_{p}\right)} + \underbrace{U\left(f,p^{m}\left(v_{p}\right)\right)-U\left(X^{f},\gamma_{0}^{f},m\right)}_{\text{Unexplained}}$$

Using the female parameters:

$$U\left(X^{m},\gamma_{0}^{m},f\right)-U\left(X^{m},m\right) = \underbrace{U\left(m,p^{f}\left(v_{p}\right)\right)-U\left(X^{m},m\right)}_{\text{Explained by } p^{f}\left(v_{p}\right)} + \underbrace{U\left(X^{m},\gamma_{0}^{m},f\right)-U\left(m,p^{f}\left(v_{p}\right)\right)}_{\text{Unexplained}}$$

# C Life Cycle Model

In this appendix I describe a life cycle model of schooling choice, from which the specification of university attendance in Section 2 is derived. I begin with a multi-period model, and then simplify it to two periods that correspond to periods 1 and 2 in Section 2.

In this model, individuals maximize lifetime utility by choosing how many years of schooling to complete. Labour is supplied inelastically. Assuming log utility for consumption, lifetime utility given schooling level $S$ is:

$$U(S) = \sum_{t=0}^{\infty} \beta^t \ln(c_t) + g^{NP}(S) \tag{C.1}$$

Following Cameron and Taber (2004), I assume that during schooling the extent to which a student can access credit markets depends on their parents. The during-schooling rate is $R = (1-r)$. After schooling, they can borrow and lend as much as they want at the market rate, which is normalized to equal the discount rate, $\beta$. If the present value of lifetime income, net of direct costs, at schooling level $S$,is $I_S$, then a youth's lifetime budget constraint is:

$$\sum_{t=0}^{S-1} \left(\frac{1}{R}\right)^t c_t + \left(\frac{1}{R}\right)^S \sum_{t=S}^{\infty} \beta^{t-S} c_t \leq I_s(e) \tag{C.2}$$

Maximizing (C.1) subject to (C.2) yields the following optimal paths for consumption:

$$\begin{aligned} c_t &= c_0 (R\beta)^t \quad t < S \\ c_t &= c_0 (R\beta)^S \quad t \geq S \end{aligned}$$

These consumption paths can be expressed as function of lifetime income, by first expressing the budget constraint as a function of $c_0$ and then factoring out $c_0$:

$$\begin{aligned} I_S &= \sum_{t=0}^{S-1} \left(\frac{1}{R}\right)^t c_t + \left(\frac{1}{R}\right)^S \sum_{t=S}^{\infty} \beta^{t-S} c_t \\ &= \sum_{t=0}^{S-1} \left(\frac{1}{R}\right)^t c_0 (R\beta)^t + \left(\frac{1}{R}\right)^S \sum_{t=S}^{\infty} \beta^{t-S} c_0 (R\beta)^S \\ &= \sum_{t=0}^{S-1} \beta^t c_0 + \sum_{t=S}^{\infty} \beta^t c_0 \\ &= c_0 \sum_{t=0}^{\infty} \beta^t \\ c_0 &= I_S (1-\beta) \end{aligned}$$

Then, this expression is substituted into the optimal consumption paths, which are in turn substituted into the utility function to derive the value of schooling:

$$V_s = \sum_{t=0}^{S-1} \beta^t \ln\left(I_S (1-\beta)(R\beta)^t\right) + \sum_{t=S}^{\infty} \beta^t \ln\left(I_S (1-\beta)(R\beta)^S\right) + g^{NP}(S)$$

To further specify $I_S$, I assume that labour is supplied inelastically until period $T$.[30] I also assume there are no earnings during schooling and direct costs of schooling, including tuition, supplies, and fees, is $DC_{t+1}$ Under these assumptions, the present value of income, for schooling level $S$ is:

$$\begin{aligned} I_S &= \left(\frac{1}{R}\right)^S \sum_{t=S}^{T} \beta^{t-S} w_{ts} - \sum_{t=0}^{S-1} \left(\frac{1}{R}\right)^t DC_{t+1} \qquad \text{(C.3)} \\ &= \left(\frac{1}{R}\right)^S W_S - \sum_{t=0}^{S-1} \left(\frac{1}{R}\right)^t DC_{t+1} \end{aligned}$$

Since I examine only the decision to attend university, I assume that there are only two levels of schooling, $S = \{0, 1\}$, which compares enrolling in university to the next best alternative, which I assume involves fewer years of schooling. With only two level, the value of schooling simplifies to:

$$\begin{aligned} V_1 &= \frac{\ln(1-\beta)}{1-\beta} + \frac{1}{1-\beta}\ln(I_1) + \frac{\ln(R\beta)}{1-\beta} + g^{NP}(1) \\ V_0 &= \frac{\ln(1-\beta)}{1-\beta} + \frac{1}{1-\beta}\ln(I_0) + g^{NP}(0) \end{aligned}$$

With only two levels of schooling, and assuming there are no direct costs when $S = 0$, lifetime income simplifies to:

$$\begin{aligned} I_1 &= \left(\frac{1}{R}\right) W_1 - DC_1 \\ I_0 &= W_0 \end{aligned}$$

The decision rule can now be summarized as attend university if $V_1 - V_0 > 0$, otherwise do not, where the net value of attending university is:

$$\begin{aligned} V_1 - V_0 &= \frac{1}{1-\beta}\left[\ln(I_1) - \ln(I_1)\right] + \frac{\ln(R\beta)}{1-\beta} + g^{NP}(1) + g^{NP}(1) - g^{NP}(0) \\ &= \frac{1}{1-\beta}\left[\ln\left(\left(\frac{1}{R}\right) W_1 - DC_1\right) - \ln(W_0)\right] + \frac{\ln(R\beta)}{1-\beta} + g^{NP}(1) + g^{NP}(1) - g^{NP}(0) \end{aligned}$$

This expression is a non-linear function of university earnings, direct costs and the rate of borrowing during university. To derive the linear specification in Section 2, I use a linear

[30] There are, of course, important gender differences in labour supply throughout the life-cycle, and these differences can be captured by differences in lifetime wages.

approximation of $\ln I_1$ evaluated where $I_1 = I_0$. Specifically, where $DC_1 = 0$, $W_1 = W_0$ and $R = 1$.

That linear approximation is:

$$V_1 - V_0 \approx \text{constant} + \underbrace{\frac{1}{1-\beta}\frac{W_1 - DC_1}{W_0}}_{g^W} + \underbrace{\frac{2}{1-\beta}}_{g^R}R + \underbrace{g^{NP}(1) - g^{NP}(0)}_{g^{NP}}$$

# D Identification

Carneiro et al. (2003) describe generally the conditions under which the distributions of a set of unobserved factors are non-parametrically identified from a system of measurement equations. In this appendix, I describe how the identification arguments from Carneiro et al. (2003) apply in my model. In order to develop the intuition, I begin with a simple one factor model, and then I show how that applies to the three factor setting. I suppress the individual and gender subscripts.

Ignoring observed covariates, suppose the latent index for dropping out is a function of a single unobserved factor, called cognitive skills:

$$U = \gamma_0 + \lambda_{0\theta_1}\theta_1 + u_0 \tag{D.1}$$

Because the intercept and the mean of the factor, $\mathbb{E}[\theta_1]$, are not separately identified, I assume that $\mathbb{E}[\theta_1] = 0$. As Carneiro et al. (2003) discuss, it is necessary to have at least two measurements for each factor. Here, I can use PISA reading scores and grades as noisy measures of $\theta_1$. The measurement errors ($u_1$ and $u_0$) are assumed to be independent of $\theta_1$ and mutually independent.

$$\begin{aligned} PISA &= \beta_{10} + \lambda_{1\theta_1}\theta_1 + u_1 \\ grades &= \beta_{20} + \lambda_{2\theta_1}\theta_1 + u_2 \end{aligned}$$

The identifying information comes from the variance-covariance matrix of the outcome and measurements. If I call the matrix of outcomes $Y$, then the variance-covariance matrix is:

$$cov(Y) = \Lambda\Sigma_\Theta\Lambda' + D_u \tag{D.2}$$

where $\Sigma_\Theta$ is a diagonal matrix of factor variances, and $D_u$ is a diagonal matrix of the measurement error variances. The elements of the diagonal in the observed matrix, $cov(Y)$, combine the measurement error and factor variances. The identification proceeds by first using the off-diagonals of $cov(Y)$ to identify the parameters of the factors and their loadings, then the diagonal of $cov(Y)$ can be used to identify $D_u$.

Because the factors have no natural scale, I normalize $\lambda_{1\theta_1} = 1$. The remaining factor parameters I need to identify are $\lambda_{0\theta_1}$, $\lambda_{2\theta_1}$ and $\sigma_{\theta_1}$. The three off diagonal elements of $cov(Y)$ are:

$$\begin{aligned} cov(U, PISA) &= \lambda_{0\theta_1}\sigma_{\theta_1} \\ cov(U, grades) &= \lambda_{0\theta_1}\lambda_{2\theta_1}\sigma_{\theta_1} \\ cov(PISA, grades) &= \lambda_{2\theta_1}\sigma_{\theta_1} \end{aligned}$$

Solving this system of three equations identifies the parameters:

$$\lambda_{0\theta_1} = \frac{cov\left(U, grades\right)}{cov\left(PISA, grades\right)}$$

$$\sigma_{\theta_1} = \frac{cov\left(U, PISA\right)}{\lambda_{0\theta_1}}$$

$$\lambda_{2\theta_1} = \frac{cov\left(PISA, grades\right)}{\sigma_{\theta_1}}$$

The factor loading $\lambda_{0\theta_1}$ is the impact that cognitive skills has on the university enrollment index. In the simple model with a single factor, the ratio of covariances that is used to identify $\lambda_{0\theta_1}$ is equivalent to using grades as an instrument for PISA.

More generally, when there are $K$ factors and $L$ equations, there are $L \times K - K$ factor loadings and $K$ variances to identify. The number off-diagonals in the covariance matrix is $L\left(1 - L\right)/2$. The order condition for identification can be summarized as $L \geq 2K + 1$. In other words, there must be at least two measures for each factor plus an additional equation. Further restrictions on the covariances are required to guarantee that the parameters can be recovered. Carneiro et al. (2003) recommend the use of dedicated measurements. I follow their suggestion using the normalization outlined in footnote 18, which requires that there is one measure that is function of only one factor. It is also necessary to have at least three equations that are a function of all of the factors.

My system is over-identified. I include 8 equations, listed below:

$$\begin{aligned}
M_0 \equiv Uni &= \lambda_{0\theta_1}\theta_1 + \lambda_{0\theta_2}\theta_2 + \lambda_{0v_p}v_p + u_0 \\
M_1 \equiv yasp &= \lambda_{1\theta_1}\theta_1 + \lambda_{1\theta_2}\theta_2 + \lambda_{1v_p}v_p + u_1 \\
M_2 \equiv PISA &= \theta_1 + u_2 \\
M_3 \equiv hmwrk &= \theta_2 + \lambda_{3v_p}v_p + u_3 \\
M_4 \equiv parasp &= \lambda_{4\theta_1}\theta_1 + \lambda_{4\theta_2}\theta_2 + v_p + u_4 \\
M_5 \equiv grades &= \lambda_{5\theta_1}\theta_1 + \lambda_{5\theta_2}\theta_2 + \lambda_{5v_p}v_p + u_5 \\
M_6 \equiv getby &= \lambda_{6\theta_2}\theta_2 + \lambda_{6v_p}v_p + u_6 \\
M_7 \equiv saved &= \lambda_{7\theta_1}\theta_1 + \lambda_{7\theta_2}\theta_2 + \lambda_{7v_p}v_p + u_7
\end{aligned}$$

The restriction that the PISA equation is a function of only $\theta_1$ is the key normalization, along with the restrictions that $\theta_1$ does not enter the *hmwrk* and *getby* equations. There are 26 non-zero covariances, which are:

$$
\begin{aligned}
cov\left(M_0,M_1\right) &= \lambda_{0\theta_1}\lambda_{1\theta_1}\sigma^2_{\theta_1}+ \quad \lambda_{0\theta_2}\lambda_{1\theta_2}\sigma^2_{\theta_2}+\lambda_{0v_p}\lambda_{1v_p}\sigma^2_{v_p}\\
cov\left(M_0,M_2\right) &= \lambda_{0\theta_1}\sigma^2_{\theta_1}\\
cov\left(M_0,M_3\right) &= \lambda_{0\theta_2}\sigma^2_{\theta_2}+\lambda_{0v_p}\lambda_{3v_p}\sigma^2_{v_p}\\
cov\left(M_0,M_4\right) &= \lambda_{0\theta_1}\lambda_{4\theta_1}\sigma^2_{\theta_1}+ \quad \lambda_{0\theta_2}\lambda_{4\theta_2}\sigma^2_{\theta_2}+\lambda_{0v_p}\sigma^2_{v_p}\\
cov\left(M_0,M_5\right) &= \lambda_{0\theta_1}\lambda_{5\theta_1}\sigma^2_{\theta_1}+ \quad \lambda_{0\theta_2}\lambda_{5\theta_2}\sigma^2_{\theta_2}+\lambda_{0v_p}\lambda_{5v_p}\sigma^2_{v_p}\\
cov\left(M_0,M_6\right) &= \lambda_{0\theta_2}\lambda_{6\theta_2}\sigma^2_{\theta_2}+\lambda_{0v_p}\lambda_{6v_p}\sigma^2_{v_p}\\
cov\left(M_0,M_7\right) &= \lambda_{0\theta_1}\lambda_{7\theta_1}\sigma^2_{\theta_1}+ \quad \lambda_{0\theta_2}\lambda_{7\theta_2}\sigma^2_{\theta_2}+\lambda_{0v_p}\lambda_{7v_p}\sigma^2_{v_p}\\
cov\left(M_1,M_2\right) &= \lambda_{1\theta_1}\sigma^2_{\theta_1}\\
cov\left(M_1,M_3\right) &= \lambda_{1\theta_2}\sigma^2_{\theta_2}+\lambda_{1v_p}\lambda_{3v_p}\sigma^2_{v_p}\\
cov\left(M_1,M_4\right) &= \lambda_{1\theta_1}\lambda_{4\theta_1}\sigma^2_{\theta_1}+ \quad \lambda_{1\theta_2}\lambda_{4\theta_2}\sigma^2_{\theta_2}+\lambda_{1v_p}\sigma^2_{v_p}\\
cov\left(M_1,M_5\right) &= \lambda_{1\theta_1}\lambda_{5\theta_1}\sigma^2_{\theta_1}+ \quad \lambda_{1\theta_2}\lambda_{5\theta_2}\sigma^2_{\theta_2}+\lambda_{1v_p}\lambda_{5v_p}\sigma^2_{v_p}\\
cov\left(M_1,M_6\right) &= \lambda_{1\theta_2}\lambda_{6\theta_2}\sigma^2_{\theta_2}+\lambda_{1v_p}\lambda_{6v_p}\sigma^2_{v_p}\\
cov\left(M_1,M_7\right) &= \lambda_{1\theta_1}\lambda_{7\theta_1}\sigma^2_{\theta_1}+ \quad \lambda_{1\theta_2}\lambda_{7\theta_2}\sigma^2_{\theta_2}+\lambda_{1v_p}\lambda_{7v_p}\sigma^2_{v_p}\\
cov\left(M_2,M_3\right) &= 0\\
cov\left(M_2,M_4\right) &= \lambda_{4\theta_1}\sigma^2_{\theta_1}\\
cov\left(M_2,M_5\right) &= \lambda_{5\theta_1}\sigma^2_{\theta_1}\\
cov\left(M_2,M_6\right) &= 0\\
cov\left(M_2,M_7\right) &= \lambda_{7\theta_1}\sigma^2_{\theta_1}\\
cov\left(M_3,M_4\right) &= \lambda_{4\theta_2}\sigma^2_{\theta_2}+\lambda_{3v_p}\sigma^2_{v_p}\\
cov\left(M_3,M_5\right) &= \lambda_{5\theta_2}\sigma^2_{\theta_2}+\lambda_{3v_p}\lambda_{5v_p}\sigma^2_{v_p}\\
cov\left(M_3,M_6\right) &= \lambda_{6\theta_2}\sigma^2_{\theta_2}+\lambda_{3v_p}\lambda_{6v_p}\sigma^2_{v_p}\\
cov\left(M_3,M_7\right) &= \lambda_{7\theta_2}\sigma^2_{\theta_2}+\lambda_{3v_p}\lambda_{7v_p}\sigma^2_{v_p}\\
cov\left(M_4,M_5\right) &= \lambda_{4\theta_1}\lambda_{5\theta_1}\sigma^2_{\theta_1}+ \quad \lambda_{4\theta_2}\lambda_{5\theta_2}\sigma^2_{\theta_2}+\lambda_{5v_p}\sigma^2_{v_p}\\
cov\left(M_4,M_6\right) &= \lambda_{4\theta_2}\lambda_{6\theta_2}\sigma^2_{\theta_2}+\lambda_{6v_p}\sigma^2_{v_p}\\
cov\left(M_4,M_7\right) &= \lambda_{4\theta_1}\lambda_{7\theta_1}\sigma^2_{\theta_1}+ \quad \lambda_{4\theta_2}\lambda_{7\theta_2}\sigma^2_{\theta_2}+\lambda_{7v_p}\sigma^2_{v_p}\\
cov\left(M_5,M_6\right) &= \lambda_{5\theta_2}\lambda_{6\theta_2}\sigma^2_{\theta_2}+\lambda_{5v_p}\lambda_{6v_p}\sigma^2_{v_p}\\
cov\left(M_5,M_7\right) &= \lambda_{5\theta_1}\lambda_{7\theta_1}\sigma^2_{\theta_1}+ \quad \lambda_{5\theta_2}\lambda_{7\theta_2}\sigma^2_{\theta_2}+\lambda_{5v_p}\lambda_{7v_p}\sigma^2_{v_p}\\
cov\left(M_6,M_7\right) &= \lambda_{6\theta_2}\lambda_{7\theta_2}\sigma^2_{\theta_2}+\lambda_{6v_p}\lambda_{7v_p}\sigma^2_{v_p}
\end{aligned}
$$

Using this system of equations, I need to solve for 20 parameters $\{\lambda_{0\theta_1},\ \lambda_{1\theta_1},\ \lambda_{4\theta_1},\ \lambda_{5\theta_1},\ \lambda_{7\theta_1}\}$, $\{\lambda_{0\theta_2},\ \lambda_{1\theta_2},\ \lambda_{4\theta_2},\ \lambda_{5\theta_2},\ \lambda_{6\theta_2},\ \lambda_{7\theta_2}\}$, $\left\{\lambda_{0v_p},\ \lambda_{1v_p},\ \lambda_{3v_p},\ \lambda_{5v_p},\ \lambda_{6v_p},\ \lambda_{7v_p}\right\}$,and $\left\{\sigma^2_{\theta_1},\ \sigma^2_{\theta_2},\ \sigma^2_{v_p}\right\}$. Solving the system proceeds in two steps. First, I express the parameters of the $\theta_1$ and $v_p$ distributions as a function of the parameters of the $\theta_2$ distribution.

$$\lambda_{5v_p} = \frac{cov\left(M_0, M_2\right)\left[cov\left(M_5, M_7\right) - \lambda_{5\theta_2}\lambda_{7\theta_2}\sigma^2_{\theta_2}\right] - cov\left(M_2, M_7\right)\left[cov\left(M_0, M_5\right) - \lambda_{5\theta_2}\lambda_{0\theta_2}\sigma^2_{\theta_2}\right]}{cov\left(M_0, M_2\right)\left[cov\left(M_4, M_7\right) - \lambda_{4\theta_2}\lambda_{7\theta_2}\sigma^2_{\theta_2}\right] - cov\left(M_2, M_7\right)\left[cov\left(M_0, M_4\right) - \lambda_{4\theta_2}\lambda_{0\theta_2}\sigma^2_{\theta_2}\right]}$$

$$\lambda_{7\theta_1} = \frac{cov\left(M_5, M_7\right) - \lambda_{5\theta_2}\lambda_{7\theta_2}\sigma^2_{\theta_2} - \left[cov\left(M_4, M_7\right) - \lambda_{4\theta_2}\lambda_{7\theta_2}\sigma^2_{\theta_2}\right]\lambda_{5v_p}}{cov\left(M_2, M_5\right) - cov\left(M_2, M_4\right)\lambda_{5v_p}}$$

$$\lambda_{0\theta_1} = \frac{cov\left(M_0, M_2\right)\lambda_{7\theta_1}}{cov\left(M_2, M_7\right)}$$

$$\lambda_{2\theta_1} = \frac{cov\left(M_1, M_2\right)\lambda_{7\theta_1}}{cov\left(M_2, M_7\right)}$$

$$\lambda_{4\theta_1} = \frac{cov\left(M_2, M_4\right)\lambda_{7\theta_1}}{cov\left(M_2, M_7\right)}$$

$$\lambda_{5\theta_1} = \frac{cov\left(M_2, M_5\right)\lambda_{7\theta_1}}{cov\left(M_2, M_7\right)}$$

$$\sigma^2_{v_p} = \frac{cov\left(M_4, M_5\right) - \lambda_{4\theta_1}\lambda_{5\theta_1}\sigma^2_{\theta_1} - \lambda_{4\theta_2}\lambda_{5\theta_2}\sigma^2_{\theta_2}}{\lambda_{5v_p}}$$

$$\lambda_{0v_p} = \frac{cov\left(M_0, M_4\right) - \lambda_{0\theta_1}\lambda_{4\theta_1} - \lambda_{0\theta_2}\lambda_{4\theta_2}}{\sigma^2_{v_p}}$$

$$\lambda_{1v_p} = \frac{cov\left(M_1, M_4\right) - \lambda_{1\theta_1}\lambda_{4\theta_1} - \lambda_{1\theta_2}\lambda_{4\theta_2}}{\sigma^2_{v_p}}$$

$$\lambda_{3v_p} = \frac{cov\left(M_3, M_4\right) - \lambda_{3\theta_2}\lambda_{4\theta_2}}{\sigma^2_{v_p}}$$

$$\lambda_{6v_p} = \frac{cov\left(M_4, M_6\right) - \lambda_{6\theta_2}\lambda_{4\theta_2}}{\sigma^2_{v_p}}$$

$$\lambda_{7v_p} = \frac{cov\left(M_4, M_7\right) - \lambda_{7\theta_1}\lambda_{4\theta_1} - \lambda_{7\theta_2}\lambda_{4\theta_2}}{\sigma^2_{v_p}}$$

In the second stage, I can express the parameters of the $\theta_2$ distribution as a function of the $\theta_1$ and $v_p$ distributions.

$$\begin{aligned}
\lambda_{6\theta_2} &= \frac{cov\left(M_6, M_7\right) - \lambda_{6v_p}\lambda_{7v_p}\sigma^2_{v_p}}{cov\left(M_3, M_7\right) - \lambda_{3v_p}\lambda_{7v_p}\sigma^2_{v_p}} \\
\sigma^2_{\theta_2} &= \frac{cov\left(M_3, M_6\right) - \lambda_{3v_p}\lambda_{6v_p}\sigma^2_{v_p}}{\lambda_{6\theta_2}} \\
\lambda_{0\theta_2} &= \frac{cov\left(M_0, M_3\right) - \lambda_{0v_p}\lambda_{3v_p}\sigma^2_{v_p}}{\sigma^2_{\theta_2}} \\
\lambda_{4\theta_2} &= \frac{cov\left(M_3, M_4\right) - \lambda_{3v_p}\sigma^2_{v_p}}{\sigma^2_{\theta_2}} \\
\lambda_{5\theta_2} &= \frac{cov\left(M_3, M_5\right) - \lambda_{3v_p}\lambda_{5v_p}\sigma^2_{v_p}}{\sigma^2_{\theta_2}} \\
\lambda_{7\theta_2} &= \frac{cov\left(M_3, M_7\right) - \lambda_{3v_p}\lambda_{7v_p}\sigma^2_{v_p}}{\sigma^2_{\theta_2}}
\end{aligned}$$

Further simplification of the first set of equations leads to quadratic functions of the parameters. Under the assumption that all of the factor loadings are positive the system can be solved. Furthermore, the solution only exists if $cov\left(M_0, M_2\right)\left[cov\left(M_4, M_7\right) - \lambda_{4\theta_2}\lambda_{7\theta_2}\sigma^2_{\theta_2}\right] - cov\left(M_2, M_7\right)\left[cov\left(M_0, M_4\right) - \lambda_{4\theta_2}\lambda_{0\theta_2}\sigma^2_{\theta_2}\right] \neq 0$ and $cov\left(M_2, M_5\right) - cov\left(M_2, M_4\right)\lambda_{5v_p} \neq 0$.