The Implication of Peer and Parental Influences on University Attendance: A Gender Comparison*

Louis N. Christofides
Michael Hoy
Joniada Milla†
Thanasis Stengos

January 4, 2012

Abstract

In this study, we explore the effect of peers and family on University attendance and graduation. We find that parental expectations and peer effects have a significant impact on the educational outcomes which operates through the interconnectedness between grades and aspirations during high school. Apart from this indirect path, parents and peers influence directly educational outcomes. Policy measures that exploit especially the parental influence on the child may be undertaken to balance the gender gap in University graduates in Canada.

Keywords: Peer and Parental Influences, University Attendance and Graduation
JEL: I20, J00

*The empirical work of this paper was carried out using Statistics Canada microdata at the University of Waterloo South-Western Ontario Research Data Center (SWORDC) and the results have been approved for release. While the research and analysis are based on data from Statistics Canada, the opinions expressed do not represent the views of Statistics Canada.

†Corresponding author. Department of Economics, University of Guelph, Guelph, Ontario, N1G 2W1, Canada. e-mail: jmilla@uoguelph.ca

Aknowledgements

We thank for their comments and suggestions the participants of the 2011 23^{rd} EALE Conference, 2010 Socio-economic conference of Statistics Canada, the 44^{th} and 45^{th} Annual Meeting of the CEA as well as participants of the October 2010 RDC Conference on Economic Relations of Parents and Children.

Louis N. Christofides
Department of Economics and Finance
Universities of Cyprus and Guelph
Kallipoleos 75, P.O. Box 20537
Nicosia 20537, Cyprus
louis.christofides@ucy.ac.cy
homepages.econ.ucy.ac.cy/~eclouis

Michael Hoy
Department of Economics and Finance
University of Guelph
Guelph, ON
N1G 2W1, Canada
519-824-4120 ext. 52169
mhoy@uoguelph.ca
www.uoguelph.ca/~mhoy

Joniada Milla
Department of Economics and Finance
University of Guelph
Guelph, ON
N1G 2W1, Canada
519-824-4120 ext. 56539
jmilla@uoguelph.ca
www.uoguelph.ca/~jmilla

Thanasis Stengos
Department of Economics and Finance
University of Guelph
Guelph, ON
N1G 2W1, Canada
519-824-4120 ext. 53917
tstengos@uoguelph.ca
www.uoguelph.ca/~tstengos

Contents

List of Tables

1 Introduction

Research in the area of education economics is driven mainly by policy oriented questions. Finding ways of motivating young people to choose education beyond high school is one of main concerns. The rationale for encouraging educational attainment is to increase the level of human capital of an economy. This later macroeconomic goal, based on growth theory, is important in explaining a significant portion of the economic growth differential between nations because a highly skilled and educated labor force assimilates new innovations and technology spillovers faster. So, higher education brings higher income in the state as well as at the individual level. Also, a more educated population produces more positive externalities for everyone in terms of lower crime rates and also would result in a pleasant social environment to live in. This is the first motivation to this paper.

The second issue is that from 1977 to 2005, the relative likelihood of females to males attending University in Canada has increased from approximately even to a ratio of 3:2. Additionally, our data indicate that in 2008, 34.55% of females and only 20.82% of males had graduated from a Bachelor's program by the age of 23. This shows that the gender gap in University degree attainment persists and is equal to 13.7 percentage points (see Table 5). From a policy point of view, the government often expresses an interest to find the policy tools that would increase the likelihood of post-secondary education attendance among the young population and balance the gender gap. This makes it necessary to identify the determinants of education decision-making among youth in Canada; how strongly are they present at different stages of high school education and do any of them contribute to the gender gap in University attendance?

The most common type of empirical study regarding the determinants of University attendance uses data that allow one to link the post secondary education (PSE) decisions of children to their cognitive and non-cognitive ability, other own characteristics, as well as to the characteristics of their family (Christofides et al., 2008; Day, 2009; Frenette, 2009). Other studies find that the differences in these variables may actually explain the gender differential. Jacob (2002) finds that 90% of the gender gap in US University attendance is due to the combined effects of differences in non-cognitive abilities of males and females

and the University premium relative to high school. Among non-cognitive abilities, he includes middle school grades and hours of homework to reflect achievement and effort, respectively, as well as two composite measures reflecting the student's compliance in the classroom and social maturity. Frenette and Zeman (2007), similar in spirit to Jacob (2002), use the YITS-A dataset and find that about 59% of the Canadian gender gap in University attendance is explained by the differences in overall marks and standardized reading tests, study habits, and progress in time through high school. Christofides et al. (2008) suggest that the resulting gender gap has its roots in the way different genders hold and revise their aspirations to attend University during high school. They find that females have higher aspirations to start with. Female students are also more likely to revise their aspirations upwards when compared to male students, producing in this way an even wider gender gap.

Our results extend those of Christofides et al. (2008). Figure 1 below shows the percentage of the students that aspire to go to University at age 15 and 17, as well as the University attendance rates at the age of 19. The black bars indicate a positive response and the grey bars a negative response. This bar chart shows how the age 17 aspirations are updated conditional on age 15 aspirations, and how this process is reflected in University attendance. Out of the 69% of females that have aspirations to attend University at the age of 15, 80% of these maintain the same aspirations at the age of 17, most of which (75%) end up attending University at the age of 19. Whereas in the case of male students aged 17, 75% of those who had University aspirations at 15 have kept the same response, but out of this group only 70% actually attend University. So, for the group of students who had University aspirations both at the age of 15 and 17, the fraction of females attending University at 19 years of age is 5 percentage points higher than males. At the age of 17, 44% of the females updated their aspirations upwards, but only 31% of the males did so. On the other hand, 20% of the female students updated their aspirations downward at the age of 17, whereas the corresponding number for males is 5 percentage points higher. Hence, female students not only are more likely to revise their aspirations upwards but also they are less likely to revise their aspirations downwards. The same is true when these aspirations are translated into University attendance at the age of 19. From the female group that updated positively their aspirations (44%) at the age of 17,

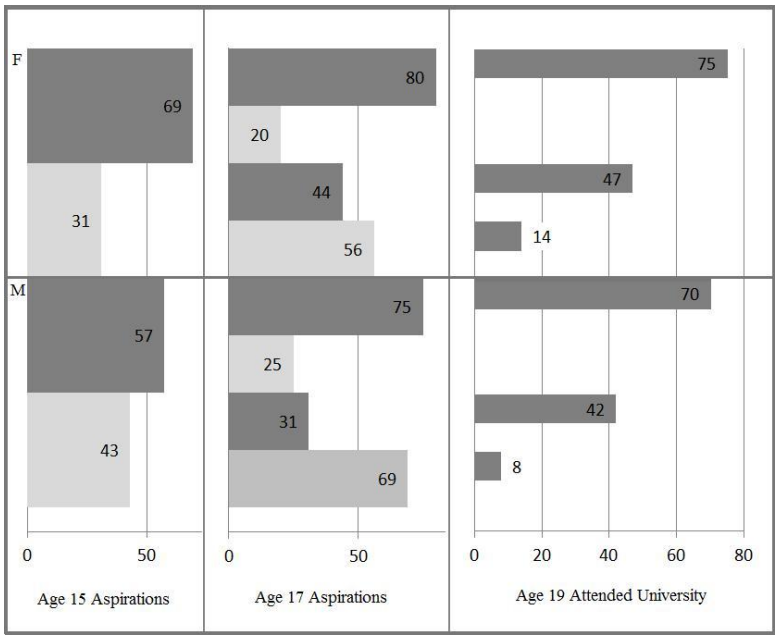

Figure 1: Percentage Aspiring for University degree and Percentage Attending

47% of them actually attended University by the age of 19. The corresponding number for the male group is 42%. Even of the students that never aspired to go to University, females are still 6 percentage points more likely to attend University than males. All of these changes contribute to a widening of the gender gap[1].

Most studies find that parental education and income contribute significantly to explaining PSE decisions. Some studies that use Canadian data are Christofides et al. (2001), Zhao et al. (2003), Johnson and Rahman (2005), Christofides et al. (2008), and Christofides et al. (2010). Some others find that parental education actually has a significant influence over and above family income (Knighton and Mirza, 2002; Finnie and Mueller, 2008). In addition to the above, we use a variable that explicitly indicates the self reported expectations that the parents have for their child related to his/her education attainment.

Apart from own individual and parental characteristics, the literature provides evi-

[1]In Figure 1 we used as the outcome the University attendance rates at the age of 19. The results remain similar when considering University attendance rates of the same students when they become 21 and 23 years old. The gender differential is also reflected in the final decision to graduate from the University program they have attended. Charts and numbers for these other three outcomes are available on request.

dence that friends in school, in the neighbourhood and in the work environment influence behaviour and decision-making of the individual including the decision to drop out of high school (see Foley et al., 2009), the decision to pursue PSE education, as well as decision-making in many social contexts such as the decision to smoke, drink alcohol, take illegal drugs or commit other crimes, and engage in safe or unsafe sexual practices. The role of peer effects was first discussed in Coleman et al. (1966) and, since then, has grown into a vast research area. A strand of this literature, relevant to the present paper, analyses peer effects on academic achievement of students, which is generally measured by standardized test scores or grade point average. Hoxby (2000a), Sacerdote (2001), Lin (2010) find important influences of peers on students' grade point average (GPA). Zimmerman (2003), Kramarz et al. (2008), Ammermueller and Pischke (2009), Boucher et al. (2010) find statistically significant though small peer effects on standardized test scores. Hanushek and Woessman (2007) question the results they obtain by stating that the causal effect of peer variables remain ambiguous and Vigdor and Nechyba (2005) find that there is no causal relationship of peer influences on academic performance. Thus, the literature has clearly not yet reached an agreement either on whether peer effects reflect a causal relationship, or on whether they are statistically significant and important in magnitude. This lack of consensus is attributed to the obstacles faced when trying to isolate and then estimate peer effects. The identification challenge in peer effect studies is due to the self-selection and the reflection problem as defined in Manski (1993). How one goes about disentangling such methodological issues is also driven by data availability. In our study we have a variety of quality characteristics of children that we can access in order to study peer effects. The YITS-A data provides us with self-reported information about each child's friends (e.g., whether a child states that "my friends smoke" or "my friends think it is OK to work hard at school"). Although the answers to such questions have the potential to capture important attitudinal characteristics of peers, finding an identification strategy to avoid the reflection and selection problem is difficult. A discussion on how we isolate peer effects is provided in section 3.

It is clear from the literature that the socio-economic status of a student and parental income and education is an important factor in decision making. We also know that differences in overall grades explain some of the gender gap in University attendance

(Frenette and Zeman, 2007; Jacob, 2002). What we do not know is what creates the difference in grades in the first place? We think that the response lies in part in the idea that motivation (aspiration to attend University) and accomplishment (academic achievement) operate as a self-reinforcing mechanism. Jacob (2002) and Frenette and Zeman (2007) attempt to explain the gender gap in University participation by academic achievement and individual characteristics. But they do not consider the possibility that academic performance and aspirations to attend University may be two simultaneous factors influencing each other and then influencing the decision to attend University. This issue is mentioned as a potential problem in Christofides et al. (2008, pp.111) but not investigated further. We think that the so-far-not-addressed interconnectedness between grades and aspirations may be one of the important reasons leading to the gender gap in academic performance (grades) and aspirations to attend University of Canadian youth. The longitudinal nature of the data allows us to probe further into the issues associated with the conditional aspiration updating as displayed in Figure 1 above.

Thus, building on the research line of Christofides et al. (2008), we try to purge the simultaneity existing between the forces determining grades and aspirations by use of instrumental variables. We construct different sets of variables for peer effects, parental expectations and outcome variables related to University attendance and University degree attainment. We track the academic performance and the aspirations of Canadian high school students to attend University using a rich Canadian dataset. Our data source, the Youth in Transition Survey - Cohort A, is a biennial longitudinal survey of 5 cycles. It follows the students involved from age 15 to age 23. We seek to analyse the role of parental influences and peer effects in determining high school outcomes for children and the formation of their aspirations about further education. We then investigate how these formed aspirations affect University attendance. Different from existing research in this area, which generally examines the issue based on a point in time and with data from a single institution (school or University), we are able to conduct a longitudinal analysis with data from several schools in Canada and thus account for the "historical" (Hanushek et al., 2003) factor in the aspirations updating process and decision making that leads to the University attendance outcome. Also, as do Black et al. (2010), we analyze the peer effects of these children while teenagers, which is an age when they are most affected by

their friends. Additionally, we are not limited to test peer effects only to high school GPA, rather we may see their effects on the students' made decisions related to attendance to University and degree completion.

Our findings suggest that during high school students are affected by both "closest friends" and/or classmates as well as by parental expectations. More specifically, a 1% increase in the mean of the overall academic level of the classmates increases the probability to achieve a higher than 70% GPA during high school by about 0.25. Self-reported peer variables, which take discrete values 0 to 3 thus four categories, potentially better measure peer influences on students. These have a persistent and mostly additive effect[2] on the outcome even though of a smaller magnitude that varies from 0.02 to 0.05 when moving up one category from the sample mean of the peer variables[3]. Parental expectations affect females' academic achievement at the age of 15, but do not have an impact on males' academic achievement until age 17 (a year prior to high school graduation). The probability of aspiring to go to University is affected by both peers and parents for female students but only by parents for male students. The probability of actually attending University after graduating high school is affected both by parental expectations and peer effects above and beyond the effect these variables have on the evolution of high school overall GPA and on the evolution of aspirations during the high school years. Even though the peer variables marginal effects are relatively small (about 0.05 to 0.10 per one category increase from the sample mean), the marginal effects from the parental expectations are substantial; the change in probability varies between 0.10 and 0.16 from a discrete change in this indicator variable from zero to one. The magnitude changes when the sample is divided by family income quartiles. It remains within the same range for females in middle and low income families but increases to 0.27 for males from middle and low income families. Peer marginal effects increase to 0.313 for females and 0.515 for

[2]By an additive or additional effect we mean a marginal effect in addition to the effect that these peer variables have on a student's overall high school GPA and aspirations about attending university during the high school years, which themselves have strong effects on the decision to attend PSE.

[3]One might argue that due to the fact that these two sets of peer variables are by construction different (identification issues arise for each), including both sets simultaneously into the regressions might be driving our results. We repeated the analysis by including each set of peer variables separately and the results are quantitatively and qualitatively similar to the ones presented in the paper. These tables are available on request.

males belonging to the middle income group. Finally, regarding the final outcome, which is the probability to graduate from University, we observe a gender difference in the effect of family environment and parental expectations.

From the perspective of a policy goal to increase PSE attendance, it would appear that a strong effect could be created by exploiting the influence that parents have on children by providing information about the advantages of PSE not only to students but also to their parents by expert counselling through the school system. Based on our results, the policy measures should focus on children of middle and low income families. This is because it is likely that the impact will be larger and because this group has a (twice as) higher gender gap in both University attendance and graduation than the high income group. Of course, it may be difficult to target by family income for a given school. But additional resources for such a program could be made available for schools in lower income districts.

The paper is organized as follows. Having introduced the topic and reviewed the existing literature in this section, we discuss the data and methodology in Sections 2 and 3, respectively. We analyze our empirical results in Section 4 and conclude in Section 5.

2 Data

The factors that influence a student's decision to pursue PSE may be grouped as financial and non-financial. While there are many papers concentrating on the financial factors (Johnson, 2008; Carmichael and Finnie, 2008; Day, 2008; Frenette, 2008), the focus in the present paper is on the non-financial factors keeping only one representative indicator from the financial factors set, namely parental income.

The source of the dataset we use is the Youth in Transition Survey - Cohort A (YITS-A), a biennial longitudinal survey of 5 waves. It follows the students involved from age 15 to age 23, from year 2000 to 2008 with interviews taking place in the spring of every two years of the time span indicated (see Table 4). In the first wave, students as well as their parents and school principals were interviewed. The first cycle of this dataset merges with the survey of OECD Programme for International Student Assessment (PISA). Beginning with the second cycle only students are interviewed. The definitions of the variables we use

in this empirical work and their availability in each cycle are provided in detail in Tables 1, 2 and 3 in the Appendix[4]. Some of the advantages of this dataset are that it contains a variety of vital information about the students, it is large in number of observations and representative of Canadian youth of that age. Also, the attrition problem from cycle to cycle is corrected by weighting every observation.

We group the variables as students own characteristics, peer characteristics, family background and parental characteristics, school and teacher characteristics and finally some regional dummy variables. The students own characteristic variables include students own aspirations for University[5] attendance, high school overall grade average for the first three cycles, the Programme for International Student Assessment (PISA) reading score, hours spent on working on homework in free time outside of school and an indicator of whether the student has information on the degree needed for a future job.

The PISA reading score is considered a reasonable proxy for cognitive abilities (Foley et al., 2009) having controlled for GPA (Jacob, 2002; Frenette, 2009). We use reading test scores of PISA, rather than math or science test scores, because the number of students writing it is higher by about fifty percent than the math and science PISA scores[6].

The self reported hours spent on homework in free time outside of school reflects the student's effort. This variable is positively correlated with grades, so we think that more time spent on homework serves as a signal indicating the diligence of the child and how persistent he/she is in the study routine rather than the other side of the coin (that the student might be struggling with the material at school).

The variables that indicate school and teacher characteristics are school size, percent of female students in the school, teacher's quality of English, and lastly the government-independent private, which is a dummy variable that equals one if the school is controlled and managed by a non-governmental organization. The aim of having this variable consists in testing whether students that attend these schools do better in terms of academic

[4]Firstly, we constructed most of the control variables as binary, taking a value one in case of success and zero otherwise, however the coefficient of determination (Pseudo-R^2) was very low. After we reconstructed them as categorical variables instead, Pseudo-R^2 almost doubled.

[5]In the Canadian educational system a college is comparable to the US junior colleges. Any University undergraduate program is four years.

[6]Additionally, Frenette and Zeman (2007) provide evidence that the PISA reading score has a higher correlation with University participation than the math and science scores.

achievement and expectations for their future education path (Day, 2009).

The peer effect variables are split into those self-reported by the student and the averaged variables. The self-reported variables are the following. "Friends smoke" is a variable that might be indicating a symptom of more general social attitudes. A teenager of age 15 that has made smoking a habit may be more likely to show negativity towards school and/or reflect an overall rebellious attitude. "Friends think it reasonable to work hard at school" is a variable capturing the fact that good students may face some negative behaviour from their classmates such as being called a "nerd" or "teacher's pet" (Cooley, 2007; Foley et al., 2009). "Friends think completing high school is important" and "Friends have plans for further education" are self reported variables used as indicators of peer pressure or general aspirations of the close group of friends. The classic way of constructing a peer variable is using the mean of the characteristics and/or outcomes of the group of friends excluding the reference student (Ammermueller and Pischke, 2009; Vigdor and Nechyba, 2005; Lee, 2007) and can be calculated by the following formula:

$$\bar{Y}_{-i} = \sum_{j=1, i \neq j}^{k} \frac{Y_j}{k} \tag{1}$$

where k is the number of the friends in the peer group. In our case this is the number of classmates excluding student i. We construct three variables using the above equation. First, Y_j takes the values of the overall high school GPA as we construct the "Grade-level average GPA". This peer variable is the proportion of students within a high school, within a grade-level having a GPA higher than 70%. It is intended to capture the influence of the general academic level of the classmates on the student i. Second, Y_j takes the values of "Aspirations to attend University", and thus using equation 1 we construct "Grade-level average aspirations". We use the same method to calculate "Grade-level average PISA" where Y_j takes the values of the PISA score of the classmates in a specific school[7].

While both represent overlapping social groups in our setting, the averaged peer variables are different from the self-reported variables. The peer group for the latter is uniden-

[7]YITS-A provides data on the students of same cohort who are at the same age but go to different high schools. So, for each school we have data only for one grade-level. Within a school, about 95% of the students participating in the questionnaire go to the same grade-level.

tified, whereas we have a defined peer group for the former, i.e. we know which student belongs to which school and grade-level. The self reported peer variables are perceptions of the student's behaviour or thoughts of his/her close friends which might not be his classmates at the same time. They are used to grasp the peer pressure that the student feels to be affected by the most. Even though they are a good measure for this purpose, the fact that we can not identify who the "closest friends" are, makes identification difficult. Regarding the averaged peer variables, in high school children typically have different classmates in each course/class and so a purer measure of classmate peer effect is not possible. However, the "quality" of children in the same year of schooling is closer than using the "quality" of children in the entire school.

The literature uses parental income and education as proxies on how parents affect their children on academic achievement and education (either through nature or nurture). Naturally, the ability of parents to help finance their children's PSE is a plausible reason why low family income may be a barrier to PSE. However, this is not likely to be the sole reason for the observed relationship between family income and PSE attendance. Income levels could also reflect many other indirect influences. For example, higher-income families may spend more on the nurturing of children in ways that allow them to develop better cognitive and non-cognitive skills related to successful entry into PSE. This process starts in childhood and is more intensive during teenage-hood. Even though parents may guide their children well and provide the necessary financial support, it is important that the child has the needed motivation to exploit these opportunities. Another indirect influence might be that the general social environment differs, on average, across income classes. Also, family income may be a signal of innate ability that is inherited by children.

In our dataset, we have further information about parents than just their income and education status. Other relevant variables are parental expectations on the educational attainment of the child and also variables indicating how these parental expectations are perceived by the child. Unlike peer effect variables considered in this paper, family represents a non-overlapping social group when it comes to parental expectations. This feature helps identification. Parental expectations take the value of 1 if parents expect the child to attain at least one University degree. A child's view of parental expectations is an indicator of how much importance the child puts on the belief of the parent; as well, it

might also be a reflection of the pressure that the parent puts on the child and the extent to which it is accepted by the child. It is often argued whether the parental expectation is a reflection of the ability of the child or just the desire of the parent for his/her child's education. We are inclined to support the interpretation that the parental expectations variable reflects the ambition of the parent for the child rather than the child's ability. In making this statement we rely on a question asked to the parent right after the question about his/her expectations on the educational attainment of the child. The question is: "What is the main reason you hope child will get this level of education" and among the responses, 68.6% of the parents responded "Better job opportunities or pay" and "Valuable for personal growth and learning" and only 9.8% responded that "Best match with child's ability".

We also include in the regressions the rural versus city dummy variable to capture the lack of opportunities and perhaps the lower level of ambition to achieve a University education for communities in the rural areas. Several dummy variables by region of province are included as well in our regressions even though they are not reported due to space constraints.

3 Methodology

The main interest in this paper is the differential in the students' educational outcomes as a consequence of their choices and, most importantly, the factors that counted more in their surrounding environment that guided the students towards these choices. In cycles 3 and 4 of our data (age 19 and 21) at least 50% of the students have already started a PSE program at a University (see Table 5). In cycle 5, most of the students are about to finish or have finished an undergraduate program in a University. For the last three cycles of the data we estimate the following equation:

$$out_{isc} = a_{isc-1}\alpha_1 + g_{isc-1}\alpha_2 + X_{isc}\gamma_1 + Z^a_{isc}\gamma_2 + P_{isc}\beta_1 + F_{isc}\beta_2 + S_{sc}\beta_3 + \varepsilon^{out}_{isc} \qquad (2)$$

where subscript i indicates the individual, s the school and c indicates the cycle. out_{isc} stands for the outcome variable, g_{ic} is the indicator for the student having an overall grade

higher than 70% and a_{isc} is the indicator of having aspirations to attend University. X_{isc} and Z^a_{isc} each is a matrix of the student's own characteristics, P_{isc} the peer variables, F_{isc} the family characteristics and parental expectations, and S_{sc} includes the school and teacher characteristics. ε^{out}_{isc} is a N(0,1) error term. The out_{ic} variable pertains to cycle 3, 4 and 5 when the students are 19, 21 and 23 years old, respectively. We use two definitions of outcome. The first is the binary variable of whether the student has Attended University and we see how variables of interest affect this outcome as students get older from 19 to 23 years old. The second outcome variable that we use is an indicator of whether the student has graduated University by the age of 23. Each of these outcomes are specifically defined in Table 1. The aim of this specification is to identify if peer pressure and parental influences during high school have any effect beyond grades and aspirations on actual University attendance and University degree completion which naturally happens after the students graduate from high school.

As mentioned earlier in section 1, there are two major issues related to peer effect estimation [8]. The first issue is the endogenous peer-group selection and also the reason why it is often argued that many empirical studies find implausibly large peer effects. For example, one would expect that a child selects her friends based on her own (unobservable) characteristics, like ability. These unobserved characteristics have a direct influence on her performance in school (grades) and this influence would be captured by the measured characteristics of her peers, thus leading to bias. Sometimes an endogeneity problem in measuring peer effects is less direct but still a substantial concern. For example, although children themselves presumably do not choose where to live and hence what school to attend, their parents' location decision may be driven to a large extent by the school district they wish to be. Unobserved parental characteristics that are reflected in this decision may well translate into unobserved characteristics or influences on their children through parental inputs. So, a child's school peers are likely to be correlated with individual unobserved characteristics of the child. Most studies point out and try to address this problem by using either classroom experimental data designed purposely to avoid self-selection into peer groups (Eisenkopf, 2010) or, when this is not possible, semi-experimental techniques. The semi-experimental techniques include the use of prox-

[8] See Nechyba (2006) for a review of the literature on peer effects identification.

ies for unobservable characteristics of children and their parents, the use of variations within schools, between grades and classes (Hoxby, 2000a; Ammermueller and Pischke, 2009) or between cohorts (Hanushek et al., 2003), and lastly the instrumental variable approach and data from natural experiments (Zimmerman, 2003; Sacerdote, 2001; Han and Li, 2009; Carrell et al., 2009).

The second problem related to the identification of peer effects, is what Manski (1993) refers to as the reflection problem - differentiating between the simultaneity of the impact of peer group on the individual and the effect of the individual on the peer group. This is one of the main econometric problems that affects estimates of the peer effects. Kramarz et al. (2008) address this by adding into the regressions a common school-grade-year effect for all students. Vigdor and Nechyba (2005) use teacher and school fixed effects to minimize the bias. Cooley (2007) states that unless student achievement is used in its lagged form, the reflection problem cannot be avoided. Hanushek et al. (2003) discuss how the lag of averaged variables instead of current values may overcome this problem only partly, and the estimates of the peer effects then might be a lower bound value. DePaola and Scoppa (2010, pp.19) use predetermined variables for ability in order to overcome the reflection problem. Day (2009) uses the YITS-A Canadian dataset. In order to minimize the bias on parameter estimates caused by this endogeneity issue, the author chooses to include a large set of explanatory variables.

Another issue closely related to the reflection problem is the separation of the endogenous (behavioural) from the exogenous (contextual) peer effects. Manski (1993) raised this concern and argued that if we can not distinguish between these two peer effects, we will not be able to answer the question of whether intervening to change one student's outcome will generate a social multiplier effect. In contrast, recognizing that peers' characteristics are difficult to distinguish from their behavior (since they are unobservable), DePaola and Scoppa (2010) follow Sacerdote (2001) and Ammermueller and Pischke (2009) and estimate a coefficient for the total peer effect. They claim that this is all that it is needed from a policy point of view.

Next we discuss how we address the identification problem arising from self-selection and reflection problem. Students self select into schools based on their own characteristics or aided by parents who consider their child's characteristics and make decisions for them.

Some of these characteristics are observable and others not (say, ability). Obviously, as an unobservable variable can not be included in regressions it will create an omitted variable bias. Moreover, classes would represent non-random samples. However, class placement as a preference of parents or students themselves is not common in high school. To mitigate further the problem of sample selection in our estimates, we control for a set of variables reflecting observable characteristics of both parents and children which also helped resolve the sample selection in Day (2009); Hanushek et al. (2003); Ding and Lehrer (2007). Among these variables we treat the PISA score as a proxy variable for students ability; parental income, education and expectations as indicator of the students socio-economic background. Additionally, the set of variables on school and teacher characteristics at least play the role of school fixed effects, which is the prevalent strategy used to address self-selection. In fact, we think that variables such as "School size", "Teacher high quality English", and especially "Government-independent Private" help further the identification related to the endogenous group selection because they provide the parents with important signals regarding the quality level of the high schools while they choose one for their child. Hence, conditional on the most important characteristics on which self-selection arises, classrooms are likely to be constructed randomly. Also even if you consider the choice of parents to live in a certain neighbourhood (that will indirectly determine the high school the child will attend), it is the age of the child that determines the grade-year he/she will enter and consequently his/her classmates. Thus, as in Friesen and Krauth (2010) we think that in this setting it is plausible to assume that parents choose schools but assignment to grade-level within school happens randomly based on the age of the child.

Having said the above, the reflection problem also is not as serious an issue using the YITS-A dataset thanks to the rich set of characteristic variables available. The coefficient estimate of "Grade-level average GPA", which is the commonly used form of behavioural peer variable, might be easily identified in the presence of "Grade-level average PISA", a pure measure of the contextual peer effects of the students in the same grade level. So, being able to control for the contextual peer effect, the coefficient of the behavioural peer variable is easily identified. Also another reasoning that helps identification is that as the group of friends increases in number (group of students is defined as the students in the same grade-level for the averaged variables), the student's individual effect on the group

decreases but the peer effect of the group on the student increases. We construct a similar variable in the aspirations equation, namely "Grade-level average aspirations", which after controlling for "Grade-level average PISA" of the other students in the same grade-level, is a measure of behavioural peer effect also. We can not make the same argument with confidence when it comes to the self-reported peer effect variables included in our analysis due to an unidentifiable peer group. Depending on the definition that the student uses to specify the closest friend group, it might constitute of only classmates or friends from the neighbourhood, or both.

For the outcome specification (equation 2), in particular, the identification problem arising from the peer variables is not present. Given that most of the students in this time of their life have started a University program, they are likely away from home and have new classmates and new friendships, hence a new environment. Meanwhile, all peer variables included in the regressions are regarding the "closest friends" and classmates during high school. So, the reflection and sample selection is not an issue in this case since the current peer group of friends is different and the only beneficial effect that might have remained from the high school friends is a contextual peer effect because both the self-reported and the averaged variables enter as predetermined variables. Hence, the causality operates only one way, from high school peer variables to the individual outcome of having attended or having graduated from a University program.

So far we have talked about the channels that might affect the outcome. Referring to the Figure 3 below, the line of discussion corresponds to the red arrows (marked Eqn. 2) and the reduced form equation 2 itself.

However, we are also interested to know how aspirations evolve up to the age of 19 (i.e. the blue arrows in the above figure (marked Eqn. 3 and Eqn. 4) corresponding to equations 4 and 3 specified below. The students are evaluated based on a set of credentials for access to an undergraduate University program. Yet, one of the main requirements of Canadian Universities is the GPA threshold. Hence, a GPA higher than the threshold would make a student eligible to attend Universities but also motivate him towards this decision. So, a high GPA may raise motivation and thus aspirations to attend University. The earlier the student has this intention (shown in Figure 1, pg.3), the more willing he would be to study harder to increase his GPA so that he may enter the program

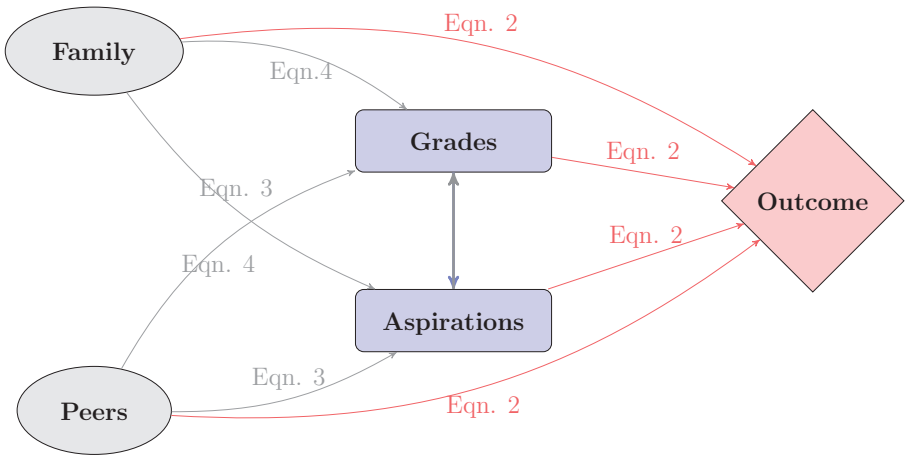

Figure 2: Model Set-up

and be accepted by the University he desires. Accordingly, a higher high school GPA achieved will induce a revision upwards of aspirations and so on. Thus, not only grades affect aspirations, but aspirations may also affect grades. Based on this idea, we have two simultaneous reduced form equations to be estimated in the first two cycles of data that coincide with the time the students are still in high school[9]. One is the academic achievement reduced form model and the other is the aspiration formation reduced form model represented as follows.

$$a_{isc} = a_{isc-1}\alpha_1^{"} + g_{isc}\alpha_2^{"} + X_{isc}\gamma_1^{"} + Z_{isc}^{a}\gamma_2^{"} + P_{isc}\beta_1^{"} + F_{isc}\beta_2^{"} + S_{sc}\beta_3^{"} + \varepsilon_{isc}^{a} \quad (3)$$

$$g_{isc} = g_{isc-1}\alpha_1^{"} + a_{isc}\alpha_2 + X_{isc}\gamma_1^{"} + Z_{isc}^{g}\gamma_2^{"} + P_{isc}\beta_1^{"} + F_{isc}\beta_2^{"} + S_{sc}\beta_3^{"} + \varepsilon_{isc}^{g} \quad (4)$$

where ε_{isc}^{g} and ε_{isc}^{a} are N(0,1) error terms and the other variables are as specified earlier. The variables included in Z_{ic}^{g} and Z_{ic}^{a} in equation 4 and 3 are the exclusion restrictions in each of the equations. Z_{ic}^{g} contains "Tutoring" and "Hours worked per week on homework" which affect aspirations only through grades. From Table 13 we can see that both these variables are highly significant in the first cycle and so are used as instruments for the grade variable in cycle 1. However, since only "Hours worked per week on homework"

[9]Regarding 2, since grades enter as predetermined variables from cycle 3, we do not have the endogeneity problem there and the outcome regressions are estimated simply by *probit*.

enters significantly in cycle 2, it will be the only IV for the grade variable in this cycle. Z_{ic}^{a} contains "Think University required for future job" and "Number of books at home" which affect grades only through aspirations. From table 14 we note that both variables have a high correlation with aspirations to attend University. As in cycle 2 only "Think University required for future job" remains significant, it will be the only instrument for the aspirations variable in cycles 2 in the instrumental variable (IV) specification. Frenette (2009) also uses YITS-A data and finds evidence that the students that know they would need a University degree for the job that they see themselves working in the future, are more likely to become enrolled in a University program. Accordingly, Table 7 displays the coefficient estimates for the first step regression of our IV estimator. Table 6 shows the correlation coefficients between the endogenous variables and the exclusion restrictions. The aspirations variable has a high correlation with "Think University required for future job" of about 0.6 and the first step regression coefficients are highly significant at 5% level. "Number of books" is also highly significant in the first step regression, although the correlation coefficient with aspirations is lower (about 0.2)[10]. The IV for grades as well enter highly significant in the first step regressions, however their correlation coefficient with the endogenous variable is not as high, especially of "Tutoring" (-0.095). Thus, as noted above our instruments are appropriate to use as they pass the relevancy test, even though the instrumental variable "Tutoring" for the Overall high school GPA might be considered a weak instrument.

The second condition to be satisfied that makes an instrument valid is the orthogonality condition of the instrument with the error term. Since this is an assumption on which the instrumental variable estimator is based on, it can not be tested. Only in the case when we have an overidentified (number of IV exceeds the number of endogenous variables) IV model we can test for overidentification[11]. Instead, we refer to the Wald test of exogeneity to show that there is reason to believe that IV is appropriate. This is a Hausman type

[10]We also considered using as an instrument an indicator variable of whether the student has received expert consulting regarding further education after graduating high school. That variable had a very low correlation (0.027) with the endogenous variable (aspirations to attend University).

[11]We do not have a overidentification test for the first cycle also, even though it is overidentified, because this test can not be performed with the maximum likelihood estimator executed by the *ivprobit* command in STATA.

test of equality between *probit* and *ivprobit*[12] specifications. For the maximum likelihood variant with a single endogenous variable, the test simply asks whether the error terms in the structural equation and the reduced-form equation for the endogenous variable are correlated[13]. If for instance the test statistic is not significant, then there is not sufficient information in the sample to reject the null that there is no endogeneity. In that case a regular *probit* regression may be appropriate; the point estimates from *ivprobit* estimation are consistent, though those from *probit* are likely to have smaller standard errors. Table 8 contains the results. In most of the cases we reject (at a 10% level the least) the null that our variable is exogenous and the Wald test provides evidence that *probit* and *ivprobit* specifications are significantly different, except for cycle 1 aspirations equation for males. One issue to take into consideration is that the exogeneity test hinges on the relevance of instruments. If the instrument is not a strong enough predictor, the standard errors are large making the test fail to reject the null hypothesis of exogeneity. This argument is a valid reason why we fail to reject the null of exogeneity of the aspirations variable in the males aspiration equation in cycle 1. As we can notice from Table 6 the correlation coefficients is low (-.072). Nevertheless, 7 the first step coefficient estimates for that regression are highly significant. Because the test results dominantly support *ivprobit* estimation rather than *probit*, we will consider the *ivprobit* results for the first two cycles so that the results are comparable.

4 Empirical Results

The empirical results are analysed differentiating between two time periods that the data covers. The first part is based on the last three cycles of data, with students of age 19 to 23, which belongs to the time period most of students have entered a PSE program and some of them have terminated schooling (graduated or dropped out). The second part involves the first two cycles of data, age 15 and 17, and it is the time period during which

[12]The STATA command *ivprobit* is an maximum likelihood instrumental variable estimator used when both the endogenous and the dependent variable are binary. This estimator computes the first step and second step coefficients simultaneously.

[13]For detailed explanation see Wooldridge(2002, pp. 472-477) "Econometric Analysis of Cross Section and Panel Data".

the students are in high school and about to graduate from high school. The data allows us to examine the outcome or the realization of the aspirations that students formed during high school in the first two cycles.

Tables 9 to 14 contain marginal effects defined as the probability change in the occurrence of the positive outcome (as indicated by the dependent variable) caused by a unit change from the mean value of the referred variable holding all independent variables at their mean levels. In the cases when the independent variable is a dummy variable the marginal effect represents the change in probability from a discrete change of the dummy variable from zero to one. In the case of the categorical variables, the marginal effects measure the impact on the probability of the positive outcome from moving one category up from the sample mean. Even though not reported in the tables because of space constrains, we control in all specifications for "Rural", "Parent(s) Immigrant", "Parental Income", "Non-birth Parent", "Percent females", "School size", "Teacher high quality English", "Government-independent Private", regional dummy variables such as "Atlantic", "Manitoba or Saskatchewan", "Alberta", "British Columbia" (see Tables 1, 2 and 3 for definitions). The asterisk attached to the marginal effect values indicate the significance level at 1% (***), 5% (**) and 10% (*).

4.1 Outcomes

We use two definitions for the outcome variable. The first involves a value of one if the student responds to have ever attended a University program and zero otherwise. The results (marginal effects) are presented in Table 9 to Table 11. The second definition of outcome involves a value of one if the student responds that he/she has graduated University and zero otherwise at the age of 23. The results (marginal effects) are presented in Table 12. Because it has been at least a year since the students graduated from high school, we do not have to worry about the simultaneity between high school grades (age 17) and outcomes (age 19, 21 and 23). Hence, we will analyse the *probit* specifications for these regressions.

Referring to Table 9, as expected, believing that the student needs a University degree for the job he is planning to work in the future, increases the probability to attend

University and to complete the program by about .20 to .30 by the age of 23. The lagged aspirations to attend University has the highest marginal effect on the probability to attend University at all ages 19, 21 and 23 among all own characteristic variables that also are significant predictors of outcomes (PISA score, high school GPA). Holding all independent variables at their means, having aspirations to attend University at the age of 17 increases the probability of attending University by 0.25. This marginal effect increases to 0.5 on the probability of a positive outcome at the age of 21, and is a little higher than that at the age of 23.

Conditional on their education, the influence of the parents seems to go beyond it through the channels of motivation or pressure. Parental expectations variable is highly significant in all the last three cycles. The estimated marginal effect stays constant at around 0.10 at the earlier age but increases as the students get older (to 0.13-0.16). The probability of attending University increases further for 19 year old male students (by 0.062) and 23 year old female students from a increase of one category above the sample mean in the importance level the student attributes to the fact that their parents think PSE is important. Having had siblings that dropped out of high school, has a negative impact on the educational outcomes of females at the age of 21 and 23. This is not true for male students of the same age[14]. The same is valid for the peer variable "Friends smoke". It is interesting to see that the effect of having had "friends that smoked" cigarettes in high school is still present (even after controlling for last year high school aspirations and overall high school GPA) in female students and has a negative marginal effect on the outcome at age 21 and 23. The additive marginal effect, i.e. beyond the effect through aspirations and overall high school GPA, is about 0.05 from an increase in one category above the mean of this variable. Another peer variable influencing probability of attendance is "Friends think continuing further education after high school". This variable, or this attitude of peers towards future education plans aspires male students positively towards University attendance by increasing their probability to attend by 0.058 at the age of 19 and by more than twice as much at the age of 23 per one category increase above the sample mean.

[14]The positive and significant marginal effect of this variable on the 19 year-old males outcome has a unexpected sign. It might be that the indirect effect of their high school sibling being a bad example encourages male students of that age to attend University.

This attitude of high school peers has a marginal effect of about 0.1 on both genders at the age of 21. Regarding the averaged peer variables, in most of the regressions they enter as insignificant. In two cases they seem to have a negative impact on the probability of the outcome and that is not what we expected[15].

We took a further step and estimated the reduced form model for three distinct quartile ranges of the parental income distribution (positive incomes considered only): lowest 25%, the interquartile range and top 25%. We did this exercise only for the last cycle of the data (when the students are 23 years old) because even those that choose to have a year off after high school by this age it is expected to be enrolled in a PSE program if they ever intended to do so. The results are displayed in Tables 10 and 11. Unlike student own characteristics, parental education looses significance when the data are decomposed by income quartile. We note that for the low and middle income group, parental expectations for University education plays an important role in increasing the probability of University attendance for both genders by about 0.11 to 0.27. Low income and high income female students' probability of the positive outcome is affected negatively by the presence of friends who smoke. In the former group this marginal effect is almost twice as big as in the latter, hinting that low income group females are more vulnerable to negative attitudes than are the females in the high income group. The middle income students seem to be affected more by positive attributes of friends like being in a class where most students aim attending University (F: 0.313; M:0.515) and having "friends who think it is ok to work hard at school" (F: 0.068) and "friends who want to continue further education after high school"(M: 0.284). For the high income family students, having a mother that has a University degree increases the probability of attending University by about 0.09 for females and 0.12 for males.

Having a look at the mean value difference between female and male students for each income category, we notice that the University attendance gap is highest for the low and middle income family students (11 and 12 percentage points, respectively) and lower for

[15]The marginal effect with the unexpected sign show up in several of our tables. Although we think it is not what we should have obtained, one explanation that we thought of is that for an increase above the mean value of the fraction of the classmates who have aspirations to attend University, the students are discouraged towards this choice. A similar interpretation we though about the average PISA score, i.e. as the percent of smarter than the average child increases in the classmates group, the reference student is discouraged towards attending University maybe feeling relatively less smarter.

high income family students (7 percentage points). So, one of the reasons why the overall gender gap seems so high in the data that are not disaggregated (10.37 percentage points, see Table 5), is because the income group that has the highest gender gap are also larger in number (about 6.6 times bigger) than the higher income group. In order to balance the gender gap, policy measures should be concentrated on the low and middle income groups.

And lastly, referring to Table 12, we can clearly see that the probability of females to graduate from a University program is increased if their parents have University education (0.143), if they expect the same for their daughters education (0.073), and if their daughter view their parents expectations as important (0.058). Meanwhile, the family environment does not have any effect on the outcome of male students at the age of 23 beyond that through overall high school GPA and aspirations about attending University held during the high school years. Notice that overall high school GPA and aspirations about attending University held during the high school years, are significant predictors of the probability to graduate University (see top panel of Table 12) for both genders apart from PISA score and "Think University required for future job". We observe that for both genders we have significant peer effects: having "Friends that smoke" decreases the probability of graduation by 0.054 in females for an increase of one category above the variable mean and "Think completing HS" increases the probability of graduating by 0.025 in males when moving up one category above the mean. By an additive or additional effect we mean a marginal effect in addition to the effect that these peer variables have on a student's overall high school GPA and aspirations about attending university during the high school years, which themselves have strong effects on the decision to attend PSE. To confirm this statement we need to answer: Would the family environment and peer effects be present not only in the final realization of the aspirations (as we have seen so far) but also during the aspiration formation process? We answer this question in the Subsections 4.2 and 4.3 below.

Black et al. (2010) and Lavy and Schlosser (2011) use the percent of females in the grade-year during high school as a measure of peer effect. They find strong evidence on the role of this variable on a set of outcomes of young adults. In our results, however, no evidence that the percent of females in high school (available only in cycle 1 of our

data[16]) has an effect on the grades or aspirations of either gender having controlled for other more precise peer variables and individual, school and parental characteristics.

4.2 Academic Achievement

In this subsection we estimate the effect of the own, peer and parental variables on the probability of achieving a GPA higher than 70%. Results of the *probit* and *ivprobit* specification for overall high school grades are shown in Table 13. For the reasons discussed in Section 3, while interpreting these results we focus mainly on the *ivprobit* specifications.

The aspirations to attend University (concurrent with overall high school GPA in this setting) enters highly significant in both age 15 and 17. We add the lagged grade variable at the age of 17 regression as in Ding and Lehrer (2007, p.304) "to pick up a variety of confounding influences including the prior, and often unrecorded as well as unobserved history of parental, school, and community effects". After we control for this variable, we notice a drastic fall in the marginal effects of the PISA score and the "Aspirations to attend University" although they remain highly significant. We see that the PISA score has a marginal effect beyond age 15 which decreases in magnitude at age 17 (for a 100 points increase in the mean PISA score the marginal effects are F: 0.09 to 0.05; M: 0.12 to 0.61) conditional on the age 15 GPA. The marginal effects of PISA tends to be lower in the IV model; nevertheless stable in magnitude around 0.10. The marginal effects of aspirations on the other hand tend to increase in the IV specifications. This increase is bigger at the age 17 regression.

Parental education significantly increases the probability of doing well in high school for females (about 0.04 at both 15 and 27 years old), but influences male students only at young age (0.061 at 15 years of age). "Parent expect University degree for child" improves the odds of doing well in school for females in cycle 1 (increases probability by 0.06). Males instead, one year prior to high school graduation (cycle 2, age 17) are negatively affected by parental expectations given that what their parent(s)' view of PSE is important to them. Female students that have siblings who dropped out of high school are affected negatively (-0.083) when 15 years old.

[16]Marginal effects not reported for this variable in our tables, however they are available on request.

Regarding peer effects, among the self-reported peer variables, "Friends who smoke" has a negative and additive (beyond the effect through age 15 overall GPA) marginal effect on high school grades at a diminishing magnitude through cycles for females (from -0.05 to -0.020 from a one category increase above the mean). For male students this peer effect is significant only at age 17, which is a critical time in their life corresponding to a year prior to high school graduation. In this time period females' probability to achieve high GPA increases with having "Friends that think completing high school is important" (0.015). Having "Friends that think of continuing further education after high school" has a positive marginal effect only in the first cycle but not beyond for both genders (F: 0.024; M: 0.031). Evaluated at the means, the peer effect increases the probability of having a high GPA at the age of 15 by about 0.25 as the proportion of students with overall GPA higher than 70% in the grade-level increases by 1% above the mean. The peer effects do not have an additive impact at the age of 17, having controlled for age 15 GPA.

4.3 Aspirations to Attend University

In this subsection we estimate the effect of the own, peer and parental variables on the probability of aspiring to go to University while the student is still in high school. Results of the *probit* and *ivprobit* specification are shown in Table 14. For the reasons discussed in Section 3, while interpreting these results we focus mainly on the *ivprobit* specifications.

Having information on the "Degree required for future job" indisputably increases the probability to aspire to attend University while in high school for both genders. The magnitude of this peer effect is substantial: at the age of 15 for females is 0.3 and 0.4 for males which decreases at the age of 17 to 0.2. Likewise in the grade equations, the marginal effect of PISA score is highly significant in the aspiration equations. We observe no pattern, however, in the magnitude of the marginal effect as the lagged dependent variable is included in the age 17 regressions. Thus, the PISA score has a significant correlation with aspirations in the *probit* regressions even though it is low in magnitude. However, the picture changes when we look at IV results. The PISA score looses significance in all regressions except at age 15 for males[17].

[17]At the age of 17 regression PISA score coefficient is also significant but in a unexpected negative sign.

The probability of aspiring to go to University during high school increases if at least one of the parents has University education. Parental expectations increase this probability for males by 0.15 at age of 15 and by half as much at the age of 17. Parental expectations increase the probability of aspiring to attend University for females by 0.048 when they are 17.

Having "friends that think completing high school is important" positively affects females' aspirations at the age of 15 and increases the probability of a positive response by 0.03 as the mean of the variable increases by one category. The "Grade-level average PISA" peer variable has a motivating effect on females at age 17 and encourages them to attend University by increasing its probability by 0.06 as the sample mean of this variable goes up by 100 points. Note that male students aspirations are not affected by peer variables: even though some of them were significant, their sign was not as expected.

5 Conclusion

In this paper we track the academic performance and the aspirations to attend University of Canadian students during high school and onwards using a rich Canadian dataset. We seek to analyse the role of parental influences and peer effects in determining high school outcomes for children and the formation of their aspirations about further education. We then investigate how these formed aspirations affect the probability to attend University and the probability to graduate from University. Different from research in this area, which generally examines the issue based on a point in time and with data from a single institution (school or University), thanks to the dataset available to us, we are able to conduct a longitudinal analysis with data from several schools in Canada. Also, like Black et al. (2010), we analyse the peer effects of these children while teenagers, which is an age when they are most vulnerable to peer pressure. Additionally, we are not limited to test peer effects only to high school GPA; rather, we may see their effects on the ultimate decision of the students related to University attendance and University degree completion.

We may conclude confidently that the PISA score is an important determinant of the academic achievement for students in high school. As expected, having good cognitive

skills is not enough to obtain a high GPA. Effort (hours worked on homework in the free time outside school) is also a very significant determinant for achieving a high GPA. Apart from these variables, peers with smoking habits is significantly negatively related to the likelihood that the student does well in school for both genders with a much bigger impact for females than for males. Regarding the students' aspirations of attending a University degree, the role of the parent seems crucial at the time of decision-making (i.e. cycle 2). Students' educational outcomes (both definitions) are significantly determined by aspirations held during the last year of high school. Parental expectations are important to both genders' University attendance outcome between ages 19-23 but important only for females when we look at the probability of degree completion.

In a nutshell, our findings suggest that during high school students are affected by both "closest friends" and/or classmates as well as by parental expectations. More specifically, a 1% increase in the mean of the overall academic level of the classmates increases the probability to achieve a higher than 70% GPA during high school by about 0.25. Self-reported peer variables, which take discrete values 0 to 3 thus four categories, are potentially better measures of peer influences on students. These have a persistent and mostly additive effect[18] on the outcome even though of a smaller magnitude that varies from 0.02 to 0.05 when moving up one category from the sample mean of the peer variables[19]. Parental expectations affect females' academic achievement at the age of 15, but do not have an impact on males' academic achievement until age 17 (a year prior to high school graduation). The probability of aspiring to go to University is affected by both peers and parents for female students but only by parents for male students. The probability of actually attending University after graduating high school is affected both by parental expectations and peer effects above and beyond the effect these variables have on the evolution of high school overall GPA and on the evolution of aspirations during

[18]By an additive or additional effect we mean a marginal effect in addition to the effect that these peer variables have on a student's overall high school GPA and aspirations about attending university during the high school years, which themselves have strong effects on the decision to attend PSE.

[19]One might argue that due to the fact that these two sets of peer variables are by construction different (identification issues arise for each), including both sets simultaneously into the regressions might be driving our results. We repeated the analysis by including each set of peer variables separately and the results are quantitatively and qualitatively similar to the ones presented in the paper. These tables are available on request.

the high school years. Even though the peer variables marginal effects are relatively small (about 0.05 to 0.10 per one category increase from the sample mean), the marginal effects from the parental expectations are substantial; the change in probability varies between 0.10 and 0.16 from a discrete change in this indicator variable from zero to one. The magnitude changes when the sample is divided by family income quartiles. It remains within the same range for females in middle and low income families but increases to 0.27 for males from middle and low income families. Peer marginal effects increase to 0.313 for females and 0.515 for males belonging to the middle income group. Finally, regarding the final outcome, which is the probability to graduate from University, we observe a gender difference in the effect of family environment and parental expectations.

From the perspective of a policy goal to increase PSE attendance, it would appear that a strong effect could be created by exploiting the influence that parents have on children by providing information about the advantages of PSE not only to students but also to their parents by expert counselling through the school system. Having parents and their children attend the same information meetings could be very productive as this would influence not only the expectations of both parents and children but reinforce the children's belief about their parents interest in possible PSE attendance. It is important that parents are aware of the difference it makes in the lifestyle (e.g. higher income) of their children if they complete a degree from a higher education post secondary institution, in our case University. In this way they will motivate and guide their children towards University education. And these students will have a peer effect on their friends and so on, creating a social multiplier effect along with the direct effect on the reference child. Based on our results, the policy measure should focus on the children of middle and low income families. This is because it is likely that the impact will be larger and this group has a (twice as) higher gender gap in both University attendance and graduation than the high income group. Of course, it may be difficult to target by family income for a given school. But additional resources for such a program could be made available for schools in lower income districts.

Appendix

Table 1: Variable Definitions

Cycles	Variable Name	Definition
1-3	Overall high school GPA	Dummy Variable. 1 if the students reports to have a high school grade point average (GPA) upto the time of interview within the range of 70-79% or higher.
1-3	Aspirations to Attend University	Dummy Variable. 1 if the highest level of education respondent think he/she will get/would like to get is a University diploma or certificate below Bachelor's, a Bachelor's Degree or higher (or one University degree or more than one university degree for cycles 1,2); 0 otherwise.
3,4,5	Attended University	Dummy Variable. 1 if response to the question Highest level of PSE taken across all programs and institutions is a University diploma or certificate below Bachelor's, Bachelors degree or higher, 0 otherwise. The universe also includes respondents who may have graduated from this level, may still be in the program or maybe left a program.
4,5	Graduated University	Dummy Variable. 1 if response to the question What is the highest degree you have attained? is a University diploma or certificate below Bachelor's, a Bachelor's degree or higher; 0 otherwise.
1	Tutoring	Categorical Variable. Equals 1 if response to the question "During the last three years, have you attended any of these special courses outside of your school to improve your results?....Private tutoring" is "Yes,sometimes" and 2 if "Yes, regularly" and 0 if "No, never".
1-3	Hours worked per week on homework	Categorical Variable. Equals 0 if "No time" spent working on homework outside class during free periods and at home within a week; 1 if "less than 1 hour a week"; 2 if "1-3 hours a week"; 5.5 if "4-7 hours a week" ; 11 if " 8-14" hours a week; 15 if "more than 15 hours a week".
1	PISA Score	Programme for International Student Assessment (PISA) reading test score expressed in per 100 points.
1,2,4,5	Think University required for future job	Dummy Variable. 1 if response to the question How much education do you think is needed for this type of work? ... One University degree?" or ... More than than one University degree?" is "Yes"; 0 otherwise. Covers respondents who have decided what type of career of work they would be interested in having when they will be about 30 years old.

Table 2: Variable Definitions (cont'd)

Cycles	Variable Name	Definition
1-3	Grade-level average grade	The portion of students in the same grade level that indicate to have an overall GPA of 70% or higher excluding the student.
1-3	Grade-level average aspirations	The portion of students in the cohort/school that indicate to have aspirations to attend University excluding the student.
1-3	Grade-level average PISA	The average PISA score of the students in the cohort that are in the same grade-level excluding the student.
1	Friends smoke	Categorical Variable. Equals 0 if student response to the question "Think about your closest friends. How many of these friends ... smoke cigarettes?" is "None of them"; 1 if "Some of them"; 2 if "Most of them"; 3 if "All of them".
1	Friends think it ok to work hard at school	Categorical Variable. Equals 0 if student response to the question "Think about your closest friends. How many of these friends ... think it's okay to work hard at school?" is "None of them"; 1 if "Some of them"; 2 if "Most of them"; 3 if "All of them".
1	Friends think completing high school is important	Categorical Variable. Equals 0 if student response to the question "Think about your closest friends. How many of these friends ... think completing high school is very important?" is "None of them"; 1 if "Some of them"; 2 if "Most of them"; 3 if "All of them".
1,2	Friends think of pursuing further education after HS	Categorical Variable. Equals 0 if student response to the question "Think about your closest friends. How many of these friends ... are planning to further their education or training after leaving high school?" is "None of them"; 1 if "Some of them"; 2 if "Most of them"; 3 if "All of them".
1	Sibling drop out	Dummy Variable. 1 if any of the child's brother's or sisters (included step, half, adopted also included siblings the child is not currently living with) is a high school drop-out; 0 otherwise.
1	Number of books at home	Categorical Variable. Equals 0 if response about the number of books at home is "None"; 1 if "1-10 books"; 2 if "11-50 books"; 3 if "51-100 books"; 4 if "101-250 books"; 5 if "more than 500 books".
1	Rural	Dummy Variable. 1 if the response to the derived variable: Indicator of rural or urban geography, based on the Statistical Area Classification, based on the 1996 Census geography equals 1 if "Rural"; 0 if "Urban".
1	Parent(s) immigrant	1 if at least one the parents has ever been a landed immigrant to Canada; 0 otherwise.
1	Parental income	variable indicating the combined (respondent and spouse/partner) total income divided by the number of the household members.
1	Non-birth parent	Dummy Variable. 1 if the parent is not by birth (i.e. by adoption, foster, step parent or guardian); 0 otherwise.
1	Father University	Dummy Variable. 1 if the father has a University certificate or diploma below Bachelor's, a Bachelor's Degree or higher; 0 otherwise.
1	Mother University	Dummy Variable. 1 if the mother has a University certificate or diploma below Bachelor's, a Bachelor's Degree or higher; 0 otherwise
1	Parent(s) expect(s) University degree for child	Dummy Variable. 1 if response of the parent to the question What is the highest level of education that you hope child will get? is "One University degree" or "More than one University degree"; 0 otherwise.

Table 3: Variable Definitions (cont'd)

Cycles	Variable Name	Definition
1	Father view of PSE important	Categorical Variable. Equals 0 if response of the child to the question How important is it to your parent(s) that you get more education after high school? To your father or other male guardian is "Not important at all", "I don't know", "No such person";1 if "Slightly important"; 2 if "Fairly important"; 3 if "Very important".
1	Mother view of PSE important	Categorical Variable. Equals 0 if response of the child to the question How important is it to your parent(s) that you get more education after high school? ... To your mother or other female guardian is "Not important at all", "I don't know", "No such person";1 if "Slightly important"; 2 if "Fairly important"; 3 if "Very important".
2	Parents view of PSE important	Categorical Variable. Equals 0 if response of the child to the question How important is it to your parent(s) that you get more education after high school? is "Not important at all", "I don't know", "No such person";1 if "Slightly important"; 2 if "Fairly important"; 3 if "Very important".
1	Percent females	This index is the ratio between the number of girls and the total enrollment (the number of boys plus number of girls)- i.e., the number of girls in the school divided by the total enrollment.
1	School size	An index represents the total enrollment in the school and is the sum of the number of boys and the number of girls enrolled in a particular school.
1	Teacher high quality English	Number of full-time teachers who have a third level qualification (i.e. a BA degree with a major in English language and literature) plus 0.5 times the number of part-time teachers with a third level qualification divided by the total number of teachers in a school.
1	Government-independent private	Dummy Variable. 1 if the school is government-independent private, 0 otherwise Government-independent private schools were coded as 1, if the school principal reported that the school was controlled and managed by a non-governmental organization (e.g., a church, a trade union or a business enterprise) or if its governing board consisted mostly of members not selected by a public agency, where it received less than 50 per cent of its core funding from government agencies.
1	Atlantic	Dummy Variable. 1 if province of the student is either of the Newfoundland, Prince Edward Island, Nova Scotia or New Brunswick; 0 otherwise.
1	Manitoba or Saskatchewan	Dummy Variable. 1 if province of the student is either Manitoba or Saskatchewan; 0 otherwise.
1	Alberta	Dummy Variable. 1 if province of the student is Alberta; 0 otherwise.
1	British Columbia	Dummy Variable. 1 if province of the student is British Columbia; 0 otherwise.

Table 4: Reference Time and Age of the Respondents by Cycle in YITS-A

Cohort A	Age	Reference Time Period	Time of the Interview
Cycle 1	15	Jan1998-Dec1999	Jan2000-Apr2000
Cycle 2	17	Jan2000-Dec2001	Jan2002-Apr2002
Cycle 3	19	Jan2002-Dec2003	Jan2004-Apr2004
Cycle 4	21	Jan2004-Dec2005	Jan2006-Apr2006
Cycle 5	23	Jan2006-Dec2007	Jan2008-Apr2008

Table 5: Gender Differential in Variables of Interest

	Aspired to Attend University			Attended University			Graduated University		
	F	M	Diff.	F	M	Diff.	F	M	Diff.
Age 15	68.68	56.62	12.06						
Obs.No.	8834	7278							
Total	12863	12849							
Age 17	68.5	56.31	12.19						
Obs.No.	8243	6736							
Total	12033	11962							
Age 19	67.53	54.03	13.5	61.59	52.41	9.18			
Obs.No.	6977	5413		4499	2936				
Total	10331	10019		7305	5602				
Age 21				66.76	56.03	10.73	4.09	2.19	1.9
Obs.No.				4820	3304		343	169	
Total				7220	5897		8388	7703	
Age 23				68.64	58.27	10.37	34.55	20.82	13.73
Obs.No.				4118	2927		2313	1284	
Total				5999	5023		6694	6168	

Table 6: Correlation Coefficients between IV and the Endogenous Variables

	Age 15		Age 17	
	F	M	F	M
IV for Aspirations				
Think University required for future job	0.565	0.576	0.530	0.574
Number of books at home	0.140	0.190	0.129	0.176
IV for Grades				
Hours worked per week on homework	0.174	0.138	0.174	0.163
Tutoring	-0.095	-0.108	-0.061	-0.061

Table 7: First Step Regression Coefficients

	Age 15		Age 17	
	F	M	F	M
IV for Aspirations				
Think University required for future job	.417***	.428***	.360***	.415***
Number of books at home	.014***	.017***		
IV for Grades				
Hours worked per week on homework	.011***	.006***	.010***	.009***
Tutoring	-.035**	-.078***		

Note: Significance levels: 0.01***, 0.05**, 0.10*.

Table 8: Wald Test of Exogeneity

	Age 15		Age 17	
	F	M	F	M
Grades Equations (p-value)	.072	.000	.022	.000
Aspiration Equations (p-value)	.013	.584	.020	.000

Table 9: Peer and Parental Influences on Probability to Attend University at the Age of 19, 21 & 23

Attended University[1]	Age 19		Age 21		Age 23	
	F	M	F	M	F	M
Think University required for future job	0.105***	0.215***	0.263***	0.330***	0.179***	0.288***
PISA score	0.043***	0.070***	0.157***	0.160***	0.097***	0.123***
Overall HS GPA	0.115***	0.141***	0.205***	0.181***	0.111*	0.101*
Lag of aspirations to attend University	0.251***	0.242***	0.483***	0.432***	0.550***	0.542***
Father University	0.085***	0.060**	0.096***	0.117**	0.040	0.076
Mother University	0.053**	0.048	0.104***	0.010	0.075**	0.094*
Parent expect Univeristy degree for child	0.104***	0.099***	0.129***	0.162***	0.126***	0.144***
Parents view of PSE important	0.026	0.062***	0.013	0.063	0.056*	0.006
Sibling HS drop out	0.014	0.111**	-0.185***	0.005	-0.170**	0.082
Friends smoke	-0.004	-0.012	-0.045**	-0.037	-0.053***	-0.007
Friends think ok to work hard	0.005	-0.010	-0.007	0.013	0.010	0.022
Friends think to complete HS	0.030*	-0.020	0.025	0.015	0.030	-0.024
Friends think continuing further education after HS	-0.021	0.058*	0.095**	0.089*	0.000	0.188***
Grade level average aspirations	-0.107	-0.332***	0.080	0.170	0.148	0.191
Grade level average PISA	-0.010	0.051	-0.046	-0.026	-0.080**	0.020
Obs.No.	4655	4240	3454	2770	2887	2390
Pseudo-R^2	0.363	0.376	0.459	0.434	0.532	0.517

[a] Significance levels: 0.01***, 0.05**, 0.10*.The table presents marginal effects evaluated at the mean. For binary variables, the estimates represent the change in probability from a discrete change from zero to one. Even though not reported in the table because of space constrains, we control for "Rural", "Parent(s) Immigrant", "Parental Income", "Non-birth Parent", "Percent females", "School size", "Teacher high quality English", "Government-independent Private", regional dummy variables such as "Atlantic","Manitoba or Saskatchewan", "Alberta", "British Columbia" (see Tables 1, 2 and 3 for definitions).

Table 10: Peer and Parental Influences on Probability to Attend University at the Age of 23 by Income Quartile - Females

Y = Females Attended University[1]	Lowest 25%		Interquartile range		Top 25%	
	dF/dx	$\bar{x}$	dF/dx	$\bar{x}$	dF/dx	$\bar{x}$
Think University required for future job	0.224***	0.568	0.157**	0.600	0.088*	0.592
PISA score	0.085***	5.544	0.081**	5.702	0.465***	0.758
Overall HS GPA	0.178**	0.893	0.169	0.920	0.103***	5.828
Lag of Aspirations to attend University	0.626***	0.650	0.543***	0.719	-0.077**	0.940
Father University	0.018	0.226	0.040	0.384	0.009	0.522
Mother University	0.005	0.195	0.051	0.295	0.089**	0.455
Parents expect University degree for child	0.173***	0.705	0.131**	0.760	0.026	0.824
Parents view of PSE important	0.049	2.789	0.047	2.851	0.127**	2.888
Sibling HS drop out	-0.140	0.067	-0.351**	0.035	-0.037	0.049
Friends smoke	-0.091***	0.805	-0.029	0.789	-0.059**	0.771
Frinds think it okay to work hard at school	0.025	2.163	0.068**	2.185	-0.029	2.265
Friends think completing HS is important	0.026	2.521	0.015	2.514	0.044	2.680
Friends think continuing further education after HS	-0.034	1.892	0.058	1.910	-0.005	1.954
Grade-level average aspirations	0.211	0.602	0.313*	0.649	-0.156	0.672
Grade-level average PISA	-0.118**	5.333	-0.110*	5.421	0.061	5.491
Obs.No.	1513		689		356	
Pseudo-R^2	0.538		0.559		0.582	
Mean Y	0.644		0.714		0.765	

[a] Significance levels: 0.01***, 0.05**, 0.10*.The table presents marginal effects evaluated at the mean. For binary variables, the estimates represent the change in probability from a discrete change from zero to one. Even though not reported in the table because of space constrains, we control for "Rural", "Parent(s) Immigrant", "Parental Income", "Non-birth Parent", "Percent females", "School size", "Teacher high quality English", "Government-independent Private", regional dummy variables such as "Atlantic","Manitoba or Saskatchewan", "Alberta", "British Columbia" (see Tables 1, 2 and 3 for definitions).

Table 11: Peer and Parental Influences on Probability to Attend University at the Age of 23 by Income Quartile - Males

Y = Males Attended University[1]	Lowest 25%		Interquartile range		Top 25%	
	dF/dx	$\bar{x}$	dF/dx	$\bar{x}$	dF/dx	$\bar{x}$
Think University required for future job	0.289***	0.404	0.321***	0.490	0.308***	0.431
PISA score	0.115***	5.344	0.162***	5.472	0.121***	5.603
Overall HS GPA	0.058	0.812	0.344***	0.868	0.194***	0.857
Lag of Aspirations to attend University	0.606***	0.548	0.495***	0.599	0.341***	0.731
Father University	0.076	0.282	0.130	0.349	-0.029	0.574
Mother University	0.016	0.218	0.062	0.347	0.116*	0.546
Parents expect University degree for child	0.113**	0.643	0.268***	0.719	0.034	0.849
Parents view of PSE important	0.061	2.721	-0.011	2.839	-0.015	2.851
Sibling HS drop out	0.200*	0.038	-0.167	0.033	0.093*	0.052
Friends smoke	-0.019	0.765	0.014	0.744	0.009	0.765
Frinds think it okay to work hard at school	0.032	1.871	0.018	1.926	-0.025	1.968
Friends think completing HS is important	0.004	2.425	-0.022	2.431	-0.061	2.476
Friends think continuing further education after HS	0.088	1.868	0.284***	1.868	0.123	l1.867
Grade-level average aspirations	-0.029	0.606	0.515**	0.649	0.443**	0.666
Grade-level average PISA	-0.045	5.383	-0.010	5.452	0.067	5.505
Obs.No.	1458		620		312	
Pseudo-R^2	0.522		0.579		0.658	
Mean Y	0.535		0.595		0.699	

[a] Significance levels: 0.01***, 0.05**, 0.10*.The table presents marginal effects evaluated at the mean. For binary variables, the estimates represent the change in probability from a discrete change from zero to one. Even though not reported in the table because of space constrains, we control for "Rural", "Parent(s) Immigrant", "Parental Income", "Non-birth Parent", "Percent females", "School size", "Teacher high quality English", "Government-independent Private", regional dummy variables such as "Atlantic","Manitoba or Saskatchewan", "Alberta", "British Columbia" (see Tables 1, 2 and 3 for definitions).

Table 12: Peer and Parental Influences on Probability to have Graduated University by the Age of 23

Graduated University[1]	Age 23			
	F	$\bar{x}$	M	$\bar{x}$
Think University required for future job	0.118***	0.559	0.092***	0.384
PISA score	0.060***	5.616	0.043***	5.347
Overall HS GPA	0.206***	0.889	0.081***	0.800
Lag of aspirations to attend University	0.321***	0.647	0.229***	0.528
Father University	0.143***	0.310	0.005	0.309
Mother University	-0.002	0.257	0.013	0.279
Parent expect University degree for child	0.073**	0.728	0.035	0.658
Parents view of PSE important	0.058**	2.798	0.011	2.721
Sibling HS drop out	-0.127***	0.054	-0.055	0.048
Friends smoke	-0.054***	0.815	-0.009	0.793
Friends think ok to work hard	0.000	2.160	-0.003	1.864
Friends think to complete HS	0.004	2.520	0.025*	2.391
Friends think continuing further education after HS	0.013	1.892	0.033	1.848
Grade level average aspirations	0.075	0.628	0.017	0.620
Grade level average PISA	0.013	5.384	-0.005	5.402
Obs.No.	3131		2832	
Pseudo-R^2	0.313		0.310	

[a] Significance levels: 0.01***, 0.05**, 0.10*.The table presents marginal effects evaluated at the mean. For binary variables, the estimates represent the change in probability from a discrete change from zero to one. Even though not reported in the table because of space constrains, we control for "Rural", "Parent(s) Immigrant", "Parental Income", "Non-birth Parent", "Percent females", "School size", "Teacher high quality English", "Government-independent Private", regional dummy variables such as "Atlantic","Manitoba or Saskatchewan", "Alberta", "British Columbia" (see Tables 1, 2 and 3 for definitions).

Table 13: Peer and Parental Influences on the Probability to Have a Overall High School GPA Higher then 70%

Overall High School GPA[1]	Age 15				Age 17			
	probit		ivprobit		probit		ivprobit	
	F	M	F	M	F	M	F	M
Tutoring	-0.053***	-0.057***	-0.055***	-0.059***	-0.008	-0.002	-0.008	-0.012
Hours of HW	0.012***	0.011***	0.013***	0.011***	0.007***	0.010***	0.007***	0.007***
PISA score	0.093***	0.149***	0.087***	0.121***	0.052***	0.077***	0.048***	0.061***
Aspirations to attend University	0.085***	0.125***	0.142***	0.336***	0.037***	0.078***	0.128***	0.276***
Lag of overall GPA					0.225***	0.354***	0.207***	0.317***
Father University	0.046**	0.063**	0.034*	0.033	0.027	0.042	0.021	0.021
Mother University	0.010	0.075***	0.014	0.061**	0.041**	-0.022	0.039**	-0.026
Parent expect Univeristy degree for child	0.065***	0.098***	0.057**	0.038	0.014	-0.009	0.003	-0.045*
Father's view of PSE important	0.000	0.036**	0.002	0.026				
Mother's view of PSE important	0.009	-0.006	-0.005	-0.021				
Parents' view of PSE important					0.015*	0.058***	0.010	0.043***
Sibling HS drop out	-0.078***	-0.062*	-0.083***	-0.053	-0.034	-0.011	-0.033	-0.019
Friends smoke	-0.052***	-0.039***	-0.050***	-0.021	-0.021***	-0.035***	-0.020***	-0.033***
Friends think ok to work hard	0.007	-0.003	0.010	0.001	0.011	-0.013	0.012	-0.017
Friends think to complete HS	-0.007	0.036**	-0.008	0.023	0.017*	-0.006	0.015*	-0.007
Friends think continuing further education after HS	0.033***	0.032*	0.024**	0.031*	-0.006	-0.013	-0.007	-0.019
Grade level average grade	0.235***	0.241***	0.245***	0.262***	-0.051	-0.046	-0.051	-0.027
Grade level average PISA	-0.088***	-0.082***	-0.072***	-0.085***	-0.027*	0.012	-0.029*	0.005
Obs.No.	6227	5721	5118	4637	5933	5515	5905	5487
Pseudo-R^2	0.256	0.225			0.293	0.245		

[a] Significance levels: 0.01***, 0.05**, 0.10*.The table presents marginal effects evaluated at the mean. For binary variables, the estimates represent the change in probability from a discrete change from zero to one. Even though not reported in the table because of space constrains, we control for "Rural", "Parent(s) Immigrant", "Parental Income", "Non-birth Parent", "Percent females", "School size", "Teacher high quality English", "Government-independent Private", regional dummy variables such as "Atlantic","Manitoba or Saskatchewan", "Alberta", "British Columbia" (see Tables 1, 2 and 3 for definitions).

Table 14: Peer and Parental Influences on the Probability to Aspire Attending University at Early Age (High School Years)

Aspirations to Attend University[1]	Age 15				Age 17			
	probit		ivprobit		probit		ivprobit	
	F	M	F	M	F	M	F	M
Think University required for future job	0.434***	0.464***	0.317***	0.429***	0.357***	0.455***	0.289***	0.217***
Number of books	0.014**	0.024***	0.005	0.018***	0.004	0.000	-0.002	-0.009
PISA score	0.050***	0.086***	-0.014	0.064**	0.037***	0.068***	-0.006	-0.036*
Overall HS GPA	0.075***	0.076***	0.723***	0.220	0.061**	0.074***	0.620***	0.705***
Lag of aspirations to attend University					0.124***	0.181***	0.083***	0.067*
Father University	0.008	0.106***	-0.004	0.078***	0.096***	0.092***	0.078***	0.014
Mother University	0.069***	0.031	0.064**	0.054**	0.042	0.034	0.032	0.037
Parent expect University degree for child	0.112***	0.165***	0.035	0.153***	0.080***	0.108***	0.048*	0.047*
Father's view of PSE important	0.042***	-0.018	0.031	0.014				
Mother's view of PSE important	0.022	0.075**	0.025	0.064***				
Parents' view of PSE important					0.044***	0.048***	0.021	-0.020
Sibling HS drop out	-0.028	-0.017	0.042	-0.006	0.014	0.035	0.049	0.050
Friends smoke	0.006	-0.029*	0.043**	-0.002	-0.027**	0.015	0.001	0.050***
Friends think ok to work hard	0.012	0.026*	0.004	0.026**	0.000	0.008	-0.007	0.007
Friends think to complete HS	0.035**	-0.029*	0.030*	-0.011	0.016	0.026*	-0.002	0.019
Friends think continuing further education after HS	-0.007	0.035*	-0.027	0.030*	0.029	0.028	0.019	0.006
Grade level average aspirations	-0.024	0.150	-0.047	-0.007	-0.112*	-0.138**	-0.107*	-0.104*
Grade level average PISA	-0.002	-0.028	0.025	0.002	0.031	0.036	0.060**	0.038
Obs.No.	5209	4755	5119	4644	6207	5871	6209	5868
Pseudo-R^2	0.414	0.421			0.304	0.394		

[a] Significance levels: 0.01***, 0.05**, 0.10*.The table presents marginal effects evaluated at the mean. For binary variables, the estimates represent the change in probability from a discrete change from zero to one. Even though not reported in the table because of space constrains, we control for "Rural", "Parent(s) Immigrant", "Parental Income", "Non-birth Parent", "Percent females", "School size", "Teacher high quality English", "Government-independent Private", regional dummy variables such as "Atlantic","Manitoba or Saskatchewan", "Alberta", "British Columbia" (see Tables 1, 2 and 3 for definitions).

References

Andreas Ammermueller and Jorn-Steffen Pischke. Peer effects in european primary schools: Evidence from progress in international reading literacy study. *Journal of Labor Economics*, 27:315–348, 2009.

Sandra Black, Paul Devreux, and Kjell Salvanes. Under pressure? the effects of peers on outcomes of young adults. Technical report, UCD Centre for Economic Research, Working Paper Series WP10/16, 2010.

Vincent Boucher, Yann Bramoulle, Habiba Djebbari, and Bernard Fortin. Do peers affect student achievement? evidence from canada using group size variation. Discussion Paper 4723, Institute for the Study of Labor, IZA, P.O. Box 7240, 53072 Bonn, Germany, January 2010.

Lorne Carmichael and Ross Finnie. Family income, access to post-secondary education and student grants: Why equal access requires more than loans. In Ross Finie, Richard E. Mueller, Arthur Sweetman, and Alex Usher, editors, *Who Goes? Who Stays? What Matters? Accessing and Persisting in Post-Secondary Education in Canada*, chapter 14, pages 347–366. Queen's School of Policy Studies, McGill - Queen's University Press, 2008.

Scott E. Carrell, Richard L. Fullerton, and James E. West. Does your cohort matter? measuring peer effects in college achievement. *Journal of Labour Economics*, 27:439–464, 2009.

Louis N. Christofides, Jim Cirello, and Michael Hoy. Family income and post-secondary education in canada. *Canadian Journal of Higher Education*, 31(1):177–208, 2001.

Louis N. Christofides, Michael Hoy, Zhi (Jane) Li, and Thanasis Stengos. The evolution of aspirations for university attendance. In Ross Finie, Richard E. Mueller, Arthur Sweetman, and Alex Usher, editors, *Who Goes? Who Stays? What Matters? Accessing and Persisting in Post-Secondary Education in Canada*, chapter 5, pages 109–134. Queen's School of Policy Studies, McGill - Queen's University Press, 2008.

Louis N. Christofides, Michael Hoy, and Ling Yang. Participation in canadian universities: The gender imbalance (1977-2005). *Economics of Education Review*, 29:400–410, 2010.

James S. Coleman, Ernest Q. Campbell, Carol J. Hobson, Alexander M. Mood James McPartland, Frederic D. Weinfeld, and Robert L. York. Equality of educational opportunity. *Washington, D.C.: U.S. Government Printing Office.*, 1966.

Jane Cooley. Alternative mechanisms of peer achievement spillovers: Implications for identification and policy. WCER Working Paper No.2008-08, Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research, November 2007.

Kathleen Day. A tangled web: The relationship between persistence and financial aid. In Ross Finie, Richard E. Mueller, Arthur Sweetman, and Alex Usher, editors, *Who Goes? Who Stays? What Matters? Accessing and Persisting in Post-Secondary Education in*

Canada, chapter 13, pages 327–346. Queen's School of Policy Studies, McGill - Queen's University Press, 2008.

Kathleen Day. The effect of high school resources on investment in post-secondary education in canada. 2009.

Marian DePaola and Vincenzo Scoppa. Peer group effects on academic performance of italian students. Working Paper Version, 2010.

Weili Ding and Steven F. Lehrer. Do peers affect student achievement in china's secondary schools? *The Review of Economics and Statistics*, 89(2):300312, May 2007.

Gerald Eisenkopf. Peer effects, motivation, and learning. *Economics of Education Review*, 29:364374, 2010.

Ross Finnie and Richard Mueller. The backgrounds of canadian youth and access to post-secondary education : New evidence from the youth in transition survey. In Ross Finie, Richard E. Mueller, Arthur Sweetman, and Alex Usher, editors, *Who Goes? Who Stays? What Matters? Accessing and Persisting in Post-Secondary Education in Canada*, chapter 4, pages 79–108. Queen's School of Policy Studies, McGill - Queen's University Press, 2008.

Kelly Foley, Giovanni Gallipoli, and David A. Green. Ability, parental valuation of education and the high school dropout decision. IFS Working Paper W09/21, Institute for Fiscal Studies, September 2009.

Mark Frenette. Why are lower-income students less likely to attend university? evidence from academic abilities, parental influences ans financial constraints. In Ross Finie, Richard E. Mueller, Arthur Sweetman, and Alex Usher, editors, *Who Goes? Who Stays? What Matters? Accessing and Persisting in Post-Secondary Education in Canada*, chapter 11, pages 347–366. Queen's School of Policy Studies, McGill - Queen's University Press, 2008.

Mark Frenette. Career goals in high school : Do students know what it takes to reach them and does it matter? Research Paper Series 11F0019M No.320, Statistics Canada, Social Analysis Division, 24-J, R.H. Coats Building, 100 Tuney's Pasture Driveway, Ottawa K1A 0T6, October 2009.

Mark Frenette and Klarka Zeman. Why are most university students women? evidence based on academic performance, study habits and parental influences. Analytical Studies-Research Paper Series 11F0019 No.303, Statistics Canada, Bussiness and Labour Market Analysis, 24-J, R.H. Coats Building, 100 Tuney's Pasture Driveway, Ottawa K1A 0T6, September 2007.

Jane Friesen and Brian Krauth. Sorting, peers, and achievement of aboriginal students in british columbia. *Canadian Journal of Economics*, 43:1273–1301, 2010.

Li Han and Tao Li. The gender difference of peer influence in higher education. *Economics of Education Review*, 28:129–134, 2009.

Eric A. Hanushek and Ludger Woessman. The role of school improvement in economic development. Working Paper 1911, CESIFO, February 2007. Category 5: Fiscal Policy, Macroeconomics and Growth.

Eric A. Hanushek, John F. Kain, Jacob M. Markman, and Steven G. Rivkin. Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18:527 44, 2003.

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

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

D. R. Johnson and F. Rahman. The role of economic factors, including the level of tuition, in individual university participation decisions in canada. *Canadian Journal of Higher Education*, 35(3):83–99, 2005.

David Johnson. How is variation in tuition across canadian provinces related to university participation in the youth in transition survey? In Ross Finie, Richard E. Mueller, Arthur Sweetman, and Alex Usher, editors, *Who Goes? Who Stays? What Matters? Accessing and Persisting in Post-Secondary Education in Canada*, chapter 12, pages 299–326. Queen's School of Policy Studies, McGill - Queen's University Press, 2008.

Tamara Knighton and Sheba Mirza. Postsecondary participation: the effects of parents' education and household income. *Education Quarterly Review*, 8(3):25–32, 2002.

F Kramarz, S Machin, and A Ouazad. What makes a test score? the respective contributions of pupils, schools, and peers in achivement in english primary education. Discussion Paper 3866, IZA, 2008.

Victor Lavy and Analia Schlosser. Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 3:1–33, 2011.

Lungfei Lee. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics*, 140(2):333–374, 2007.

Xu Lin. Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics*, 28(4):825–860, 2010. URL `http://www.journals.uchicago.edu/doi/pdf/10.1086/653506`.

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

Thomas J. Nechyba. Income and peer quality sorting in public and private schools. In Eric A. Hanushek and Finis Welch, editors, *Handbook of Economics of Education*, volume 2, chapter 22, pages 1327–1367. Elsevier B.V., 2006.

Bruce Sacerdote. Peer effects with random assignments: Results for dartmouth roommates. *The Quarterly Journal of Economics*, pages 681–704, May 2001.

Jacob Vigdor and Thomas Nechyba. Peer effects in north carolina public schools. *Unpulished*, 2005.

John Zhao, Miles Corak, and Garth Lipps. Family income and participation in postsecondary education. *Analytical Studies Branch Research Paper Series 2003210e, Statistics Canada, Analytical Studies Branch*, 2003.

David J. Zimmerman. Peer effects in academic outcomes: Evidence from a natural experiment. *The Review of Economics and Statistics*, 85(1):9–23, 2003.