The Dynamics of Entry and Exit in Post-Secondary Education

by

Joniada Milla

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctor of Philosophy
in
Economics

Guelph, Ontario, Canada

ABSTRACT

THE DYNAMICS OF ENTRY AND EXIT IN POST-SECONDARY EDUCATION

Joniada Milla
University of Guelph, 2013

Advisor:
Thanasis Stengos

This thesis brings to the forefront of the existing literature the importance of analyzing transitional dynamics among different levels of schooling and to the labor market. I perform empirical analyses using confidential longitudinal survey data from Statistics Canada; employing program evaluation techniques, and regression modelling.

The first chapter is joint work with Louis Christofides, Michael Hoy and Thanasis Stengos. We explore the forces that shape the development of aspirations and the achievement of grades during high school and the role that these aspirations, grades, and other variables play in educational outcomes such as going to university and graduating. We find that parental expectations and peer effects have a significant impact on educational outcomes through grades, aspirations, and their interconnectedness. Apart from this indirect path, parents and peers also influence educational outcomes directly. Policy measures that operate on parental influences may modify educational outcomes in desired directions.

The second chapter estimates the wage returns to university quality. I distinguish between two distinct measures of university quality. The first is a survey-based university reputation ranking, and the second is a new ranking, which I construct from several university characteristics in order to objectively reflect the university quality. The findings

indicate that the wage returns of having a Bachelor's degree from a highly ranked university are 10.3% for women, and 13.4% for men. The returns are higher when comparing the wages in the top and bottom tails of the ranking distribution and gender differences are identified.

The third chapter is a large-scale study on how students form and revise expectations. This affects their decision to drop out and/or change field of study once they have accessed post-secondary education. I find evidence that students change expectations and educational pathways as they are exposed to unexpected new information. This informs them about the quality of match between their own ability and the program that they are enrolled. Using non-parametric methods I show that this relationship is not linear.

ACKNOWLEDGEMENTS

I am very grateful to my advisor, Professor Thanasis Stengos, for his continued support in every single day of my doctoral studies. This thesis would not have been possible without his assistance and guidance. I would also like to acknowledge the invaluable help of Professor Michael Hoy, who contributed a lot more than was expected of him as a committee member. I am indebted to Professors Miana Plesca, Louis Christofides and Alex Maynard for their feedback and advise in important stages of my education and dissertation process.

I thank all the staff and faculty at the Department of Economics for a very friendly and supportive environment. In addition, I am grateful to the Statistical Analyst at SWORDC, Dr. Pat Newcombe-Welch, for being so patient and always accommodating all my data and output requirements.

My thanks also go to my friends Katerina Koka, Mehmet Pinar, Diana Alessandrini, Jennifer Teng, Fraser Summerfield, Kathleen Rodenburg and Marcel Oestreich who made this experience as enjoyable as possible.

I wish to thank my mother, father and brother - Hasime, Arben and Elton Milla - to whom I owe everything. I also thank my best friend, Sabuhi Yahyayev, for his encouragement, love and endless belief in me.

Finally, I acknowledge that while the research and analysis in this doctoral thesis are based on data from Statistics Canada, the opinions expressed do not represent the views of Statistics Canada.

Table of Contents

List of Tables

List of Figures

Chapter 1

Grades, Aspirations and Post-Secondary Education Outcomes

1.1 Introduction

The importance of education in general, and post-secondary education (PSE) in particular, in the process of individual human capital acquisition and fulfillment, the externalities that stem from PSE and the significance of this fabric of knowledge for the process of economic growth are now well accepted. Apart from the many theoretical explorations, a large empirical literature has emerged. The empirical explorations focus on a number of important issues. Studies of the determinants of university attendance link the PSE decisions of children to their cognitive and non-cognitive ability, their other characteristics, and the characteristics of their family (Day, 2009; Frenette, 2009). Others explore important gender dimensions (Jacob, 2002; Frenette and Zeman, 2007; Christofides et al., 2008). Some other studies focus on the important role of parental education and income (see Zhao et al., 2003; Johnson and Rahman, 2005; Knighton and Mirza, 2002; Finnie and Mueller, 2008, among many others). Apart from own individual and parental characteristics, the empirical literature provides evidence that friends in school, in the neighborhood, and in the work environment influence the behavior and decision-making of individuals 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 (smoking, drinking alcohol, taking illegal drugs, committing crimes, and engaging in safe or unsafe sexual practices).

Entry into PSE in general and university[1] in particular is paved through the motivation or aspiration to attend university and the achievement of a high enough grade point average (GPA), mechanisms which are subject to common but also separate forces (thus providing identification). These mechanisms influence each other but also condition ultimate outcomes such as going to university. The path to university involves aspirations and achievement through high school, requiring longitudinal data to study it. The Youth in Transition Survey - Cohort A (YITS-A), a Statistics Canada data set, provides the information needed to explore the high school experience leading up to PSE outcomes. In this paper we exploit the longitudinal nature of YITS-A and the wealth of student, parental, peer and other variables that it contains to study the web of high school aspirations, grade achievement and eventual PSE outcomes.

[1] In the Canadian educational system a college is comparable to the US junior colleges. Most university undergraduate programs are four years but it is possible to graduate with a "general", three-year degree.

We analyze the role of parental and peer variables in determining student aspirations about further education and their high school performance. We then investigate how these aspirations and grades affect university attendance. To our best knowledge, this is the first study that recognizes the simultaneity involved between aspirations and grade achievements. We use the wealth of data available in YITS-A to select instruments to deal with it. We also recognize and exploit the temporal separation between aspirations/grade achievements during high school and PSE outcomes when estimating outcome equations. Different from existing research in this area, which generally examines these issues at a point in time and often using data from a single institution (school or university), we are able to conduct a longitudinal analysis with data from several schools in Canada, thus accounting for the "historical" factor (Hanushek et al., 2003) in the aspirations updating process and decision making leading to university attendance.

The literature has yet to achieve a consensus on whether peer effects are significant and whether they are causal. How one goes about disentangling identification issues in peer effects, for any specific application, is driven in part 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. We use student-elicited characteristics of their group of peers, which potentially measure peer influences more effectively and are better proxies in terms of sampling and measurement error than the peer variables commonly used in the literature. Also, as in Black et al. (2010), we analyze the peer effects on these children while teenagers, which it is often argued to be an age when they are most affected by their friends. Additionally, we are not limited to test peer effects on GPA attainment; we are also able to study their influence on student decisions to attend university and to complete a degree.

Our findings suggest that the influence of the closest friends and/or classmates as well as parents is pervasive. Peers, but not parental expectations, affect the aspirations for university and grade performance while students are in high school. Parental expectations affect eventual outcomes, such as university attendance or completion, directly and beyond any effects they may have had at the intermediate stage on high school aspirations and grades. That is, they have direct effects on outcomes as well as indirect ones, through aspirations and grades which themselves influence outcomes. These influences are conditioned by income group and gender. Females' eventual outcomes are affected by having had peers that smoked at age 15. This effect is not statistically significant for male students. We believe that these effects are well established and net of the reflection problem, sample selection problem and correlated effects that have been identified in the literature. From the perspective of a policy goal to increase PSE attendance, it would appear that the influence that parents have on children's eventual outcomes could be enhanced by providing expert counselling about the advantages of PSE not only to students but also to their parents. Based on our results, such policy measures should focus on children of low-income families because it is likely that the impact will be larger on this group.

The paper is organized as follows. We discuss the data and methodology in Sections 1.3 and 1.2, respectively. We analyze our empirical results in Section 1.4 and conclude in Section 1.5.

1.2 Methodology

The analysis in this paper is conducted separately for two periods: the PSE years, when students are 19-23 years old (cycles 3-5 of the survey), and the high school period when students are 17 years old (cycle 2 of the survey). For the PSE years we estimate the following equation:

$$Out_{is,c} = \alpha_1 Asp_{is,c-1} + \alpha_2 GPA_{is,18} + Peers_{is,15}\beta + \gamma ParentExp_{is,15} + X_{i,c-1}\theta + \varepsilon_{is,c} \tag{1.1}$$

The outcome of student i in school s and cycle c ($Out_{is,c}$) is modelled as a function of the aspirations to attend university dating one cycle earlier or two years earlier ($Asp_{is,c-1}$), high school overall grade at age 18 ($GPA_{is,18}$), a vector of peer quality measures ($Peers_{is,15}$), parental expectations ($ParentExp_{is,15}$) both at age 15, and a vector containing a comprehensive set of predetermined control variables, $X_{i,c-1}$. $\varepsilon_{is,c}$ is a N(0,1) error term. We use two definitions of PSE outcome: (i) "Attended university" at ages 19, 21, 23 separately and (ii) "Graduated university" at age 23. Among other issues, this specification explores whether peer pressure and parental influences during high school have any effect beyond their effect on grades and aspirations on actual university attendance and university degree completion.

Students are evaluated based on a set of credentials for access to an undergraduate university program. 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. The earlier the student has this intention, the more willing he would be to study harder to increase his GPA so that he may enter the program and be accepted by the university he desires. Accordingly, if a higher GPA is achieved during high school, it will induce a revision upwards of aspirations and so on. Thus, not only might grades affect aspirations, but aspirations may also affect grades. Based on this idea, we have two simultaneous reduced form equations to be estimated at age 17 when the students are still in high school.[2] Equation (1.2) below specifies the probability to achieve a high school GPA higher than 70% at age 17 as a function of aspirations to attend university at 17 and "High school GPA", peer effects and "Parental Expectations" and a set of control variables, all at age 15.

$$GPA_{is,17} = \alpha_1^{'} Asp_{is,17} + \alpha_2^{'} GPA_{is,15} + Peers_{is,15}\beta^{'} + \gamma^{'} ParentExp_{is,15} + X_{i,15}\theta^{'} + \varepsilon_{is,17}^{'} \tag{1.2}$$

Similarly, equation (1.3) defines the probability to have aspirations to attend university.[3]

$$Asp_{is,17} = \alpha_1^{''} Asp_{is,15} + \alpha_2^{''} GPA_{is,17} + Peers_{is,15}\beta^{''} + \gamma^{''} ParentExp_{is,15} + X_{i,15}\theta^{''} + \varepsilon_{is,17}^{''} \tag{1.3}$$

We use instrumental variables to estimate the parameters of the above simultaneous equations (1.2) and (1.3). The identification of these two equations is discussed in detail in

[2]Regarding equation (1.1), since grades enter as predetermined variables from cycle 3, we do not have the endogeneity problem there and the outcome regressions are estimated by *probit*.

[3]See figure 1.2 for a flow chart describing the relations involving these three equations. For simplicity, not all influences in equations (1.1), (1.2) and (1.3) are shown.

section 1.4.

1.2.1 Peer Effects

A strand of the peer effect literature, relevant to the present paper, analyses peer effects on the academic achievement of students, which is generally measured by standardized test scores or grade point average (GPA). Hoxby (2000), Sacerdote (2001), Lin (2010) find important influences of peers on students' GPA. Zimmerman (2003), Kramarz et al. (2008), Ammermueller and Pischke (2009), Boucher et al. (2010) all 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) report no causal influence from peers on academic performance.

There are two major identification issues related to peer effect estimation discussed in the literature. 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. Students self-select into schools (or via parental decisions) based on their own characteristics, some of which are observable and others not (like ability). To mitigate the problem of sample selection in our estimates, we control for a set of variables measuring 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, the PISA score is a measure for student's cognitive ability. Parental income, education, expectations, nurturing and monitoring behaviour indices and family educational support are the most comprehensive indicators of the student's family environment and socio-economic background used in the literature. Additionally, the set of variables on school and teacher characteristics control for a very rich variety of factors related to teachers, students and educational resources that may affect the quality of the high schools. These variables 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 school they choose for their child. Hence, conditional on the most important characteristics on which self-selection arises, grade-levels are likely to be constructed randomly. Also, even if one accepted the choice by parents to move to a certain neighbourhood, it is the age of the child that determines the grade-year he will enter and consequently his classmates. Thus, as in Friesen and Krauth (2010), we think that, in this setting, it is plausible to assume that even where parents choose the school, the assignment to a grade-level within a school happens exogenously, based mostly on the age of the child. A certain amount of randomness is inherent in this process and is beyond the control of the parents and the child.

Further, in the event of self-selection, students may select into peer groups with similar unobserved characteristics that are stable, at least within the adolescence years while they go to the same school and live in the same neighbourhood. We introduce the lagged grades and aspirations variable on the right-hand side of all our regressions to wipe out these common effects that otherwise would have been captured by the peer effects variable. This is based on the discussion of Hanushek et al. (2003, see page 531) who take the first difference of the dependent variable in order to eliminate the “historical influences” but state that it is equivalent to adding the lag of the outcome in the right-hand side of the regression. In this way no restriction is imposed on its coefficient. Hence, our peer effect estimates are free of self-selection and correlated effects after having conditioned on a variety of characteristics and factors based on which self-selection arises and accounting for the unobserved correlated effects.

The second issue is 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. The reflection problem arises only if we try to estimate peer effects when the outcome of interest and the peer variable (constructed as an average of the same peer outcome) are concurrent because they may simultaneously affect each other. In our setting, the only way to avoid the simultaneity problem is to use the past values instead of the concurrent peer effect variables as in Hanushek et al. (2003). More precisely, we use the average of the PISA score at age 15 of current classmates. For the self-reported variables we also use two-year lagged values. Hanushek et al. (2003, see pg.535) state that even though this strategy will identify the peer effect coefficients, it will provide a lower bound estimate.

1.3 Data

1.3.1 Sampling Characteristics

The source of the dataset we use is the YITS-A, a biennial longitudinal survey of 5 cycles. It follows the students involved from age 15 to 23, from year 2000 to 2008 with interviews taking place in the spring of every two years of the time span indicated (see Table 1.5). In the first cycle, students as well as their parents and school principals were interviewed. The first cycle of this dataset merges with the survey of the OECD Programme for International Student Assessment (PISA). Beginning with the second cycle, the students only are interviewed. The definitions of the variables we use in this empirical work are provided in detail in Appendix A. Given that the survey initially interviews only 15-year old students, most of them (93%) are registered in the same grade level. In our subsample we have 710 high schools with eleven students per school on average. Sampling in YITS-A was conducted based on a two-stage probability sampling. In the first stage the high schools were chosen, and in the second stage the students within each school were chosen. For student population representation purposes we use probability weights in all our estimations. Since the stratum in this survey is the school, we use robust standard errors clustered by school.

1.3.2 Own, Parental and School Characteristics

In order to account for the heterogeneity across individuals, we control for a set of variables measuring a student's own characteristics, peer characteristics, family background and parental characteristics, school and teacher characteristics. The students own characteristic variables include the PISA reading score at age 15, hours spent working on homework in their free time outside of school and an indicator of whether the student reports that a university degree is needed to work in the future job where the student plans to work at the age of 30. The PISA reading score is considered a reasonable proxy for cognitive abilities (Foley et al., 2009) having controlled for the high school GPA (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 tests.

The school and teacher variables include a comprehensive set of indicators measuring different aspects of the teaching quality of high schools. These include school size, the percent of female students in the school, "Teacher quality", "Instructional time", "Quality of school physical infrastructure", "Quality of school's material educational resources", "Teacher-related factors affecting school climate", "Student-related factors affecting school climate", "Grade-level average GPA" and "Grade-level average aspirations" which play the role of fixed effects by accounting for the general socio-economic level of the students in each school and for the fact that some schools may be more generous in grading but others not; "Teachers' morale and commitment", "Teacher shortage" and lastly the variable “Government-independent private”, which accounts for differences in the academic achievement and expectations for future education of the students who attend these schools (Day, 2009).

Parents may influence their children’s academic achievement and aspirations in several ways. The ability of parents to help finance their children’s PSE is a plausible reason why a low family income may be a barrier to PSE. 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 better develop the cognitive and non-cognitive skills that condition successful entry into PSE. This process starts in childhood and continues into the teenage years. Another indirect influence might be through the general social environment which differs, on average, across income classes and the education level of parents. Parental education may also be a signal of innate ability that is inherited by children. Hence, in our specifications we control for parental education and household income per capita. In our dataset, apart from the above mentioned, we have further information about parents and the student’s family background. Other variables are “Sibling drop-out”, “Parent(s) immigrant”, “Non-birth parent”, “Parent(s) view of PSE important”, “Parents’ nurturance behaviour”, “Parents’ monitoring behavior”, “Family educational support” and “Residence region indicators”.

“Parental expectations” is an indicator variable of whether the parents expect the child to attain at least one university degree in the future. Unlike the peer effect variables considered in this paper, family is a non-overlapping social group when it comes to parental expectations. This feature helps identification. Do parental expectations reflect 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 the answers to a question in the survey 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".

Aiming to capture the influence of peers, we use several peer variables most of which are self-reported by the student and one is a grade-level averaged variable. The self-reported variables are the following. "Friends smoke" might be indicating general social attitudes. A teenager of age 15 who has made smoking a habit may be more likely to show negativity towards school and/or reflect an overall rebellious attitude. "Friends think it okay to work hard" 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 HS is important" and "Friends think going to PSE" serve as indicators of the general aspirations of the close group of friends towards education. There are both advantages and disadvantages to these variables. On one hand, we have no information about the "close group of friends". Nevertheless, it is likely that most of them are friends from the neighbourhood and thus attend the same high school as the reference student. On the other hand, being reported by the students themselves, these variables are perceptions about their closest friends and the peer pressure that the students feel to be affected by the most. This is a valuable characteristic of the elicited variables. The standard 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, among many others). We construct one variable in this way, "Grade-level average PISA" which captures the influence of the cognitive ability of the classmates on student i. In high school, children typically have different classmates in each course/class and so a purer measure of classmate peer effect is not necessarily desirable. However, the "quality" of children in the same year of schooling is closer than using the "quality" of children in the entire school.

1.4 Empirical Results

Using data from YITS-A, Figure 1.1 shows the percentage of the students who aspire to go to university at age 15 and 17 by gender. It also shows university attendance rates at the age of 19. The black bars indicate a positive response and the grey bars a negative response. This figure shows how aspirations at 17 are updated, conditional on age 15 aspirations and how this process leads up to university attendance. There are two points to take away from this chart.

Firstly, for both genders the earlier they aspire to attend university the more likely they

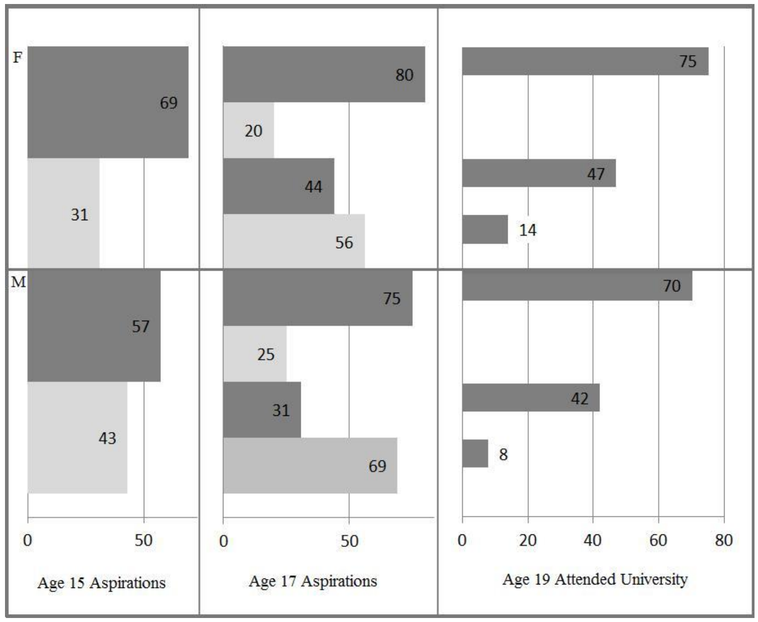


Figure 1.1: Percentage aspiring to attend and percentage that have attended a university program

are to actually attend university after graduating from high school. Secondly, male students start with lower aspirations at age 15. They are less likely to be persistent in their university aspirations, i.e. they are less likely to revise aspirations upward[4] and more likely to revise aspirations downward than females[5]. As a result they are less likely to attend university at 19 (conditional on prior aspirations) than females[6]. This information confirms that separate

[4]At the age of 17, 44% of the females who at 15 had not aspired to go to university updated their aspirations upwards, but only 31% of the males did so.

[5]On the other hand, 20% of the female students who at 15 had aspired to attend university updated their aspirations downward at 17, whereas the corresponding number for males is 25%.

[6]Out of the 69% of females that have aspirations to attend university at 15, 80% of these maintain the same aspirations at 17, most of whom (75%) end up attending university at 19. In the case of male students aged 17, 75% of those who had university aspirations at 15 keep the same response, but out of this group only 70% actually go on to university. So, for the group of students who had university aspirations both at 15 and 17, the fraction of females attending university at 19 is 5 percentage points higher than that for males. From the female group that upgraded their aspirations (44%) at 17, 47% of them actually attended university by the age of 19. The corresponding number for the male group is 42%. Even among the students that never aspired to go to university, females are 6 percentage points more likely to attend university than males.

investigations of the process leading up to PSE outcomes should be conducted for males and females.

The empirical results are analyzed differentiating between the two time periods that the data covers: the PSE years and the high school years. In Tables 1.1 to 1.4 we report 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 because of space limitations, we control in all specifications for a set of variables, which are listed in the footnote of each table. It is worthy of mentioning that among these variables, PISA score, "Parental education", "Parental income" and "Parents view of PSE important" have a positive effect on the probability of attending university. Having siblings that dropped out of high school and non-biological parents affects negatively the probability to attend university. For both genders, conditional on university attendance at age 21, students with a rural background are more likely to graduate and attain the university degree by the age of 23. Among the high school characteristics, the school size and the type of the school ("Government-independent private") have both a positive effect on the probability of PSE outcomes' realization. These variables have a similar effect on the probability to achieve a high GPA and on the probability to aspire to attend university. Other variables affecting the grades (but not PSE outcomes after controlling for high school grades) are “Teacher quality” and “Parents Nurturance Behaviour”, which have a positive influence. “Student-teaching staff ratio” and “Quality of School’s Material Educational Resources”, which measure the lack of potential time with the teacher and lack of educational resources, have a negative influence on the probability to achieve a high GPA in high school. These results all are in accordance with the findings in the literature. The full results are available on request.

1.4.1 PSE Outcomes

We investigate two outcome variables: the probability to have “Attended university” and the probability to have “Graduated university” (see Tables 1.1 and 1.2). Since it has been at least a year since students graduated from high school, simultaneity between high school grades (age 18) and outcomes (age 19, 21 and 23) is unlikely. Hence, we analyze the *probit* specifications for these regressions.

Referring to Table 1.1, the lagged “Aspiration to attend university” variable has the highest marginal effect on the probability to attend university at all ages and among all variables that also are significant predictors of outcomes. Holding all independent variables at their means, having aspirations to attend university during high school increases the probability to attend university by 0.361 for females and 0.342 for males. This marginal effect increases to 0.518 for females and 0.439 for males at age 21, and to 0.539 for females and 0.538 for males at age 23. Thus the effect of the aspirations on the actual decision is substantial.

Table 1.1: Peer and Parental Influences on Probability to Attend University at the Age of 19, 21 & 23

Attended University	Attended University						Graduated University	
	Age 19		Age 21		Age 23		Age 23	
	F	M	F	M	F	M	F	M
High School GPA	0.239***	0.316***	0.228***	0.267***	0.112*	0.069	0.157***	0.062
	(0.067)	(0.061)	(0.072)	(0.067)	(0.069)	(0.067)	(0.032)	(0.036)
	0.361***	0.342***	0.518***	0.439***	0.539***	0.538***	0.111***	0.148***
	(0.040)	(0.055)	(0.035)	(0.037)	(0.042)	(0.041)	(0.039)	(0.028)
Parental expectations	0.164***	0.254***	0.128***	0.186***	0.129***	0.167***	-0.001	0.050**
	(0.044)	(0.050)	(0.037)	(0.045)	(0.039)	(0.045)	(0.035)	(0.024)
Friends smoke	0.001	-0.029	-0.028	-0.053**	-0.064***	-0.029	-0.039**	0.002
	(0.021)	(0.031)	(0.020)	(0.027)	(0.021)	(0.029)	(0.019)	(0.016)
Friends think it okay to work hard	0.015	0.045	0.010	-0.027	0.035	-0.008	-0.004	-0.009
	(0.023)	(0.031)	(0.022)	(0.028)	(0.022)	(0.027)	(0.020)	(0.015)
Friends think completing HS	0.001	-0.011	0.021	-0.009	0.017	-0.002	0.001	0.032**
	(0.026)	(0.032)	(0.023)	(0.031)	(0.025)	(0.028)	(0.024)	(0.017)
Friends think going to PSE	0.069	0.109*	0.084	0.118	0.043	0.150***	0.027	-0.033
	(0.062)	(0.061)	(0.054)	(0.059)	(0.053)	(0.058)	(0.057)	(0.044)
Grade-level average PISA	-0.028	-0.038	-0.026	0.033	-0.055	0.011	-0.011	0.016
	(0.033)	(0.131)	(0.038)	(0.041)	(0.036)	(0.049)	(0.034)	(0.025)
Sample size	2271	1694	2816	2258	2358	1932	2258	1815
Pseudo-R^2	0.3597	0.3595	0.4760	0.4599	0.5342	0.5036	0.3697	0.3582

Note: Significance levels: 0.01***, 0.05**, 0.10*. Standard errors in parenthesis. The table presents *probit* 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 tables because of space constraints, we control in all specifications for “PISA score”, “Parent(s) immigrant”, “Parental income”, “Non-birth parent”, “Parents’ view of PSE important”, “Parents’ nurturance behaviour”, “Parents’ monitoring behaviour”, “Family Educational support”,“Sibling drop-out”,“Parental education”,“Percent females”, “School size”, “Teacher quality”, “Government-independent private”, “Instructional time”, “Quality of school physical infrastructure”, “Quality of schools’ material educational resources”, “Teacher-related factors affecting school climate”, “Student-related factors affecting school climate” ,“Teachers’ morale and commitment”, “Teacher shortage”, “Student-teaching staff ratio”, and “Residence region indicators”. In the last two columns labeled “Graduated University” we also control for “Attended University” at age 21.

The influence of parental expectations on the decision to attend university is highly significant for all ages even after conditioning on parental education, income, “Family educational support”, “Parents’ nurturance behaviour” and “Parents’ monitoring behaviour”. The estimated effect for females decreases with age from 0.164 at age 19 to 0.129 and 0.128 at age 21 and 23, respectively. The marginal effects are higher for males than females at all ages. They decrease from 0.254 at age 19, to 0.186 at 21 and to 0.167 at age 23.

It is interesting to see that the effect of “Friends smoke” (at age 15) is still present (even after controlling for past aspirations and high school overall GPA) in the outcomes of male students at age 21 and female students at age 23. Associating with friends who smoke cigarettes decreases males’ probability to attend university by 0.053 and that of females by 0.064. Another peer variable influencing the probability of attendance is “Friends think going to PSE”. This variable influences males students positively by increasing their probability to attend university by 0.109 at age 19 and by 0.150 at age 23 given a one category increase above the mean. Regarding peer cognitive ability[7], “Grade-level average PISA”, in most of the regressions is insignificant, after controlling for grades and aspirations during high school. So, it seems that after graduating from high school female students are mainly affected by parental expectations, whereas male students are affected by both parental expectations and their peers’ aspirations and attitudes.

In the last two columns of Table 1.1, we show the regression results for the outcome “Graduated university” conditional on the lag of “Attended university”. Females’ probability to graduate is negatively affected by “Friends smoke” and decreases by 0.039. On the other hand, for male students parental expectations and “Friends think completing HS is important” increase their probability to graduate by 0.050 and 0.032, respectively.

We took a further step and estimated the reduced form model for three distinct quartiles of the parental income distribution: the lowest 25%, the middle 50% and the top 25%. We

[7] 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.

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 are, by this age, be enrolled in a PSE program. The results are displayed in Table 1.2. Aspirations to attend university have an important effect for all income groups and genders. We note that, except for female students of high-income families, the parental expectations variable plays an important role in increasing the probability of university attendance for both genders. “Friends smoke” has a negative effect on the probability of the outcome of females only. This effect is nonlinear and monotonically decreasing as we go from low to middle and lastly to high-income families. The marginal effect for the lower end of the income distribution (-0.094) is almost twice as big as the one in the higher end (-0.049). This suggests that females of low-income families are more vulnerable to friends with negative attitude or rebellious behaviour, than the female students in the other income groups. Friends with aspirations to go to PSE positively affect students from high-income families. Notice that, conditional on past aspirations, only male students of high-income families are influenced by peers. Male students from low and middle-income families are only affected by parental expectations.

We note that the university attendance gap (the difference between the “Mean Y” in Table 1.2 between female and male students for each income category) is highest for the low-income family students (16.4 percentage points) and much lower for the middle (9.1) and high-income (10.6) family students. So, any attempt to balance the gender gap, should be concentrated on the low-income group students in particular.

Table 1.2: Peer and Parental Influences on Probability to Attend University at Age 23 by Income Distribution

Y = Attended University	Lowest 25%		Middle 50%		Top 25%	
	F	M	F	M	F	M
High School GPA	-0.005	0.105	0.318***	0.044	0.021	0.163*
	(0.091)	(0.116)	(0.112)	(0.084)	(0.088)	(0.103)
Aspirations to attend university	0.439***	0.538***	0.602***	0.693***	0.495***	0.483***
	(0.097)	(0.075)	(0.055)	(0.043)	(0.066)	(0.078)
Parental expectations	0.138**	0.237**	0.193***	0.138**	0.062	0.139**
	(0.067)	(0.092)	(0.060)	(0.061)	(0.052)	(0.073)
Friends smoke	-0.094***	-0.014	-0.071**	-0.062	-0.049**	-0.001
	(0.044)	(0.058)	(0.032)	(0.041)	(0.020)	(0.040)
Friends think it okay to work hard	0.025	-0.031	0.054	0.025	0.023	0.002
	(0.045)	(0.060)	(0.036)	(0.041)	(0.028)	(0.044)
Friends think completing HS	-0.005	0.045	0.044	-0.012	-0.013	-0.055
	(0.041)	(0.068)	(0.042)	(0.044)	(0.033)	(0.046)
Friends think going to PSE	0.019	0.143	0.023	0.081	0.123*	0.267***
	(0.084)	(0.143)	(0.084)	(0.087)	(0.062)	(0.098)
Grade-level average PISA	-0.065	0.033	-0.025	-0.036	-0.051	0.053
	(0.063)	(0.088)	(0.057)	(0.074)	(0.046)	(0.081)
Sample size	496	370	1203	990	701	579
Pseudo-R^2	0.549	0.525	0.609	0.563	0.535	0.5134
Mean Y	0.661	0.497	0.649	0.558	0.747	0.641
Gender gap	0.164		0.091		0.106	

Note: Significance levels: 0.01***, 0.05**, 0.10*. Standart errors in parenthesis. The table presents *probit* 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 tables because of space constraints, we control in all specifications for “PISA score”, “Parent(s) immigrant”, “Parental income”, “Non-birth parent”, “Parents’ view of PSE important”, “Parents’ nurturance behaviour”, “Parents’ monitoring behaviour”, “Family Educational support”,“Sibling drop-out”,“Parental education”,“Percent females”, “School size”, “Teacher quality”, “Government-independent private”, “Instructional time”, “Quality of school physical infrastructure”, “Quality of schools’ material educational resources”, “Teacher-related factors affecting school climate”, “Student-related factors affecting school climate”,“Teachers’ morale and commitment”, “Teacher shortage”, “Student-teaching staff ratio”, and “Residence region indicators”.

The marginal effects of the family environment and peers on the PSE outcomes can be interpreted as additional effects, i.e. a marginal effect in addition to the effect that these variables have on a student's overall high school GPA and aspirations during the high school years, which themselves have strong effects on the decision to attend university. Next, we analyze role of family and friends in the aspirations formation process during high school.

1.4.2 High School Years

In this subsection we estimate the effect of the own, peer and parental variables on the probability of achieving a GPA higher than 70% and the probability to have university aspirations at age 17, i.e. equations 1.2 and 1.3. We use instrumental variables to identify the parameters in these equations. The exclusion restriction for "High school GPA" is the change in the hours spent doing homework at home after school in the free time (ΔHours worked on HW) between age 15 and 17. The exclusion restriction for the aspirations variable is the change in the belief of whether the student thinks a university degree is required to work in the future job at age 30 (ΔThink university required for future job). We acknowledge that in the levels, these two instruments may be driven by parental influences. Nevertheless, assuming parents affect students in a more or less consistent way through time, if we take the first difference for each of these two instruments the parental effect (individual effect for each student) would be differenced out. In this way "ΔHours worked on HW" captures the change in effort as a result of an individual choice only; "ΔThink university required for future job" captures external information regarding the degree requirements or change in preferences for the future job at 30 years old.

Table 1.3: Correlation Coefficients between IV and the Endogenous Variables at age 17

	High School GPA		Aspirations to attend university	
	F	M	F	M
$\triangle$Hours worked on HW	0.062	0.069		
$\triangle$Think university required for future job			0.443	0.444

Table 1.3 shows the correlation coefficients between the instrumental variables and the endogenous variables. The first step regression estimates are presented in the bottom panel of Table 1.4. The aspirations variable has a high correlation with "ΔThink university required for future job" of 0.44 for both genders and the first step regression coefficients in Table 1.4 are significant at the 1% confidence levels. This is a relevant instrument for the aspirations variable. Because it is based on future plans, it affects the student's motivation and thus aspirations for higher education. It may affect the academic performance during high school but that may only happen through channels of aspiration formation. The probability of achieving a high GPA is positively correlated to "ΔHours worked on HW" indicating that higher achievement comes with more effort. It is also highly significant at the conventional significance levels in the first step regressions for both genders. "ΔHours worked on HW" has a direct effect on the grades and may affect aspirations indirectly only through grades.

We refer to the Wald test of exogeneity to test whether the instrumental variable estimator is different from the *probit* estimator. This is a Hausman type test of equality between the *probit* and *ivprobit*[8] specifications. For the maximum likelihood variant with a single endogenous variable, the test asks whether the error terms in the structural equation and the reduced-form equation for the endogenous variable are correlated. If the test statistic is not significant, then there is not sufficient information in the sample to reject the null that there our variable of interest is exogenous. Table 1.4 contains the results. In all of the cases we reject, at 1% and 5% significance levels, the null that our variable is exogenous. So, since the Wald test provides evidence that *probit* and *ivprobit* specifications are significantly different, we base our discussion on the results obtained by using the instrumental variable estimator for the time the students are still in high school.

The estimation results are shown in Table 1.4. The coefficients of the endogenous variables increase considerably after the instrumental variable approach is used in both grade and aspiration regressions. They are highly significant regardless of the estimator used and conditional on the measure for cognitive ability.

In both high school grades and aspirations, parental expectations have a statistically insignificant effect but the effect of peers seems to prevail for both genders. Having friends with a smoking habit decreases the probability to achieve a high school GPA higher than 70% by 0.021 for females and by 0.034 for males. An increase in classmates average cognitive ability measure increases the students' probability to do well in school by 0.047 for females and by 0.131 for males. Similar to the grades equation, female students probabil-

[8]The STATA command *ivprobit* is a maximum likelihood instrumental variable estimator used when both the endogenous and the dependent variables are binary.

Table 1.4: Peer and Parental Influences on High School Outcomes - Age 17

	High School GPA$_{17}$ $> 70\%$				*Aspirations to attend university*$_{17}$			
	probit		ivprobit		probit		ivprobit	
	F	M	F	M	F	M	F	M
Aspirations to attend university$_{17}$	0.045***	0.099***	0.256***	0.309***	–	–	–	–
	(0.015)	(0.025)	(0.078)	(0.067)				
High School GPA$_{17}$ $> 70\%$	–	–	–	–	0.081**	0.115***	0.749***	0.754***
					(0.033)	(0.030)	(0.075)	(0.007)
Parental expectations$_{15}$	0.026*	0.009	0.008	-0.011	0.126***	0.149***	0.039	0.001
	(0.017)	(0.025)	(0.020)	(0.031)	(0.027)	(0.028)	(0.054)	(0.043)
Friends smoke$_{15}$	-0.019**	-0.035***	-0.021**	-0.034**	-0.033**	0.004	0.011	0.039***
	(0.008)	(0.013)	(0.009)	(0.015)	(0.013)	(0.019)	(0.026)	(0.015)
Friends think it okay to work hard$_{15}$	0.015	-0.013	-0.016	-0.019	-0.006	0.017	-0.024	0.013
	(0.010)	(0.017)	(0.011)	(0.019)	(0.018)	(0.019)	(0.019)	(0.017)
Friends think completing HS$_{15}$	0.006	0.009	0.009	-0.009	0.016	0.011	0.011	0.008
	(0.011)	(0.017)	(0.011)	(0.018)	(0.019)	(0.019)	(0.018)	(0.018)
Friends think going to PSE$_{15}$	-0.016	0.017	-0.018	0.026	0.016	-0.007	0.017	-0.031
	(0.012)	(0.0.018)	(0.014)	(0.020)	(0.017)	(0.020)	(0.020)	(0.019)
Grade-level average PISA$_{15}$	0.058***	0.149***	0.047***	0.131***	0.084***	0.102***	0.080***	-0.009
	(0.014)	(0.029)	(0.017)	(0.030)	(0.023)	(0.026)	(0.027)	(0.035)
Sample size	4846	4438	3982	3611	4096	3772	4070	3727
Pseudo-R^2	0.3196	0.2987			0.3234	0.4197		
First Step Results								
$\triangle$Think university required for future job			0.627***	0.737***			–	–
			(0.045)	(0.048)				
$\triangle$Hours worked on HW			–	–			0.027***	0.014*
							(0.009)	(0.008)
Pseudo R^2			0.2872	0.3590			0.2433	0.1829
Wald Chi^2 Overall Regression Significance			604.80***	752.98***			453.01***	387.98***
Wald Test of Exogeneity (p-value)			0.0003	0.020			0.0472	0.0053

Note: Significance levels: 0.01***, 0.05**, 0.10*. Standard errors in parenthesis. The table presents *probit* and *ivprobit* 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 tables because of space constraints, we control in all specifications for “PISA score”, “Parent(s) immigrant”, “Parental income”, “Non-birth parent”, “Parents’ view of PSE important”, “Parents’ nurturance behaviour”, “Parents’ monitoring behaviour”, “Family Educational support”,“Sibling drop-out”,“Parental education”,“Percent females”, “School size”, “Teacher quality”, “Government-independent private”, “Instructional time”, “Quality of school physical infrastructure”, “Quality of schools’ material educational resources”, “Teacher-related factors affecting school climate”, “Student-related factors affecting school climate” ,“Teachers’ morale and commitment”, “Teacher shortage”, “Student-teaching staff ratio”, and “Residence region indicators”.

ity to have university aspirations is affected positively by the level of cognitive ability of their classmates as measured by the average PISA score. Male student's aspirations are not affected by classmates, but by the effect of the closest friends. In this regression, "Friends smoke" has a positive effect of 0.039 on the probability of aspiring to go to university for males.

1.5 Conclusion

In this paper we use a rich Canadian dataset to analyze the role of a number of variables, including parental influences and peer effects, in determining the formation of aspirations about further education and grade achievements of high school students. We then investigate how these aspirations affect the probability to attend university and the probability to complete a university degree. 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), we are able to conduct a longitudinal analysis with data representative of the Canadian youth.

We conclude that the individuals' high school GPA and the aspirations for further education held during high school are important determinants of the probability to attend university. Conditional on a measure for cognitive ability, having a GPA higher than 70% increases the probability of going to university by about 0.239 for females and 0.316 for males. Having aspirations to attend university during high school, increases the probability of attendance by 0.361 for females and 0.342 for males at age 19. At age 21 and 23 the marginal effects are higher in magnitude. The probability of attending university after graduating from high school for male students is affected both by parental expectations and peer effects above and beyond the effect that these variables have on the evolution of the overall GPA and on the evolution of aspirations during the high school years. Female students' probability to attend university is affected (through their direct channel) by parental expectations only until age 21, but by both peers and parents at 23 years old. Even though the peer variables' marginal effects are relatively small, the marginal effects from the parental expectations are substantial; the increase in probability of attendance varies with age between 0.129 and 0.164 for females and between 0.167 and 0.254 for males. When we split the sample by income group, we find that the peer effect on females' probability to attend university diminishes as we move from lower to higher income group families. In the case of males, only those from the high-income group are affected by peer aspirations to attend PSE. Regarding the high school period, after correcting for the simultaneity between grades and aspirations, we find that having aspirations to attend university is at least as important as cognitive ability in increasing the probability of high academic performance. Students are affected more by their close friends and classmates than by their parents, teachers and high school characteristics.

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. Oreopoulos and Dunn (2012) show that students from disadvantaged high schools in Toronto change significantly their aspirations regarding post-secondary education attainment after they have been exposed to an informational intervention on the returns to PSE as well as financial aid. 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 university. These students will have a peer effect on their friends during high school, creating a social multiplier effect along with the direct effect on the reference child. Based on our results, the policy measure should focus mainly on the children of low-income families because it is likely that the impact will be larger in this group. Note, also, that this group has a higher gender gap in university attendance than the middle and 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 A: Variable Definitions

Own Characteristics

High School GPA Dummy Variable. 1 if the students reports to have a high school grade point average (GPA) up to the time of interview within the range of 70-79% or higher.

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.

Aspirations to attend 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, Bachelor's degree or higher, 0 otherwise. The respondents may have graduated from this level, may still be in the program or maybe left the program.

High School GPA Dummy Variable. 1 if the students reports to have a high school grade point average (GPA) up to the time of interview within the range of 70-79% or higher.

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.

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, Bachelor's degree or higher, 0 otherwise. The respondents may have graduated from this level, may still be in the program or may be left the program.

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.

Hours worked on HW 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".

PISA score Programme for International Student Assessment (PISA) reading test score expressed in per 100 points.

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 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.

Peer Variables

Grade-level average GPA The portion of students in the same grade-level that indicate to have an overall high school grade point average (GPA) of 70% or higher excluding the reference student.

Grade-level average aspirations The portion of students in the grade-level that indicate to have aspirations to attend university excluding the student.

Grade-level average PISA The average PISA score of the students in the same grade-level excluding the reference student.

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".

Friends think it okay to work hard 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".

Friends think completing HS is important Categorical Variable. Equals 0 if student's 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".

Friends think going to PSE 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".

Family Characteristics

Sibling drop-out Dummy Variable. 1 if any of the child's brothers or sisters is a high school drop-out; 0 otherwise.

Parent(s) immigrant Equals 1 if at least one of the parents has ever been a landed immigrant to Canada; 0 otherwise.

Parental income Variable indicating the combined (respondent and spouse/partner) total income divided by the number of the household members. Total income is derived from a sum of the nine income sources collected during the parent interview. They are Wages and Salaries before deductions, including bonuses, tips and commissions; Net Income from Farm and Non-Farm Self-employment (after expense and before taxes); Employment Insurance benefits (before deduction); Canada Child Tax Benefits and provincial child tax benefits or credits (including Quebec Family Allowance); Social Assistance (welfare) and Provincial Income Supplements; Support program received, such as spousal and child support; Other Government Sources, such as Canada or Quebec Pension Plan Benefits, Old Age Security Pension, or Workers' Compensation Benefits; Goods and Service Tax Credit/Harmonized Tax Credit received in 1999; and Other Non-Government sources including dividends, interest and other investment income, employment pension, RRIFs and annuities, scholarships, and rental income.

Non-birth parent Dummy Variable. 1 if the parent is not by birth (i.e. by adoption, foster, step parent or guardian); 0 otherwise.

Parental education *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. *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.

Parental expectations 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.

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". In cycle 1 we may differentiate between mother and father.

Parents' nurturance behaviour Parent's reports on the frequency with which parents: praise child; listen to child's ideas and options; make sure child knows that they are appreciated; speak of good things those children does; and, seem proud of the things child does. A YITS scale variable is derived from this information.

Parents' monitoring behaviour Parent's reports on the frequency that the parent: know where child goes at night; know what child is doing when he/she goes out; and, know who child spends time with when he/she goes out. A YITS scale variable is derived from this information.

Family educational support Student's reports on the frequency that his/her parents and siblings work with them on their school work. A PISA index is then derived.

High School and Teacher Characteristics

Percent females This index is the ratio between the number of girls and the total enrolment (the number of boys plus number of girls), i.e. the number of girls in the school divided by the total enrolment.

School size The total enrolment in the school.

Teacher quality Number of full-time teachers who have a third level qualification (i.e. a Bachelor's 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. The third level qualifications counted are a degree in English and literature, in mathematics and science (chemistry, physics, biology or earth science).

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 percent of its core funding from government agencies.

Instructional time Instructional time for 15-year-old students in the school and derived hours of schooling per year.

Quality of school physical infrastructure School principal' reports on the extent to which learning by 15-year-olds in their school was hindered by: poor condition of buildings; poor heating and cooling and/or lighting systems; and, lack of instructional space. An index is derived from the above information.

Quality of schools' material educational resources School principals' reports on the extent to which learning by 15-year-olds in their school was hindered by: lack of instructional material; not enough computers for instruction; lack of instructional materials in the library; lack of multi-media resources for instruction; inadequate science laboratory equipment; and, inadequate facilities for the fine arts. An index of the quality of schools' educational resources is derived after.

Teacher-related factors affecting school climate Principals' reports on the extent to which the learning by 15-year-olds was hindered by: low expectation of teachers; poor student-teacher relations; teachers not meeting individual students' needs; teacher absenteeism; staff resisting change; teacher being too strict with students; and students not being encouraged to achieve their full potential. An index is derived for principals' perceptions of teacher-related factors affecting school climate.

Student-related factors affecting school climate Principals' reports on the extent to which learning by 15-year-olds in their school was hindered by: student absenteeism; disruption of classes by students; students skipping classes; students lacking respect for teachers; the use of alcohol or illegal drugs; and students intimidating or bullying other students. An index is derived for principals' perceptions of student-related factors affecting school climate.

Teachers' morale and commitment The extent to which school principals agreed with the following statements: the morale of the teachers in this school is high; teachers work with enthusiasm; teachers take pride in this school; and, teachers value academic achievement. An index was derived for principals' perceptions of teachers' morale and commitment.

Teacher shortage The principals' views on how much learning by 15-year-old students was hindered by: shortage or inadequacy of teachers in general and in each type of courses in language, math, and science. An index is derived from these information.

Student-teaching staff ratio This index is the school size divided by the total number of teachers (part-time teachers contribute 0.5 and full-time contribute 1.0).

Residence region indicators *Rural Dummy Variable*: Indicator of rural vs. urban geography, based on the Statistical Area Classification, based on the 1996 Census geography equals 1 if "Rural"; 0 if "Urban". *Atlantic Dummy Variable:* 1 if province of the student is either of the Newfoundland, Prince Edward Island, Nova Scotia or New Brunswick; 0 otherwise. *Manitoba or Saskatchewan Dummy Variable, Alberta Dummy Variable, British Columbia Dummy Variable* are constructed similarly.

Appendix B

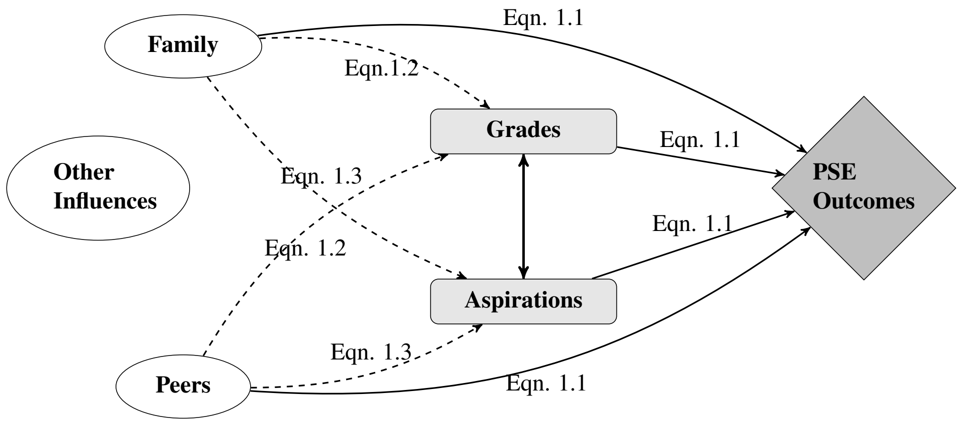


Figure 1.2: Model Set-up

Table 1.5: 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	January 1998-December 1999	January 2000-April 2000
Cycle 2	17	January 2000-December 2001	January 2002-April 2002
Cycle 3	19	January 2002-December 2003	January 2004-April 2004
Cycle 4	21	January 2004-December 2005	January 2006-April 2006
Cycle 5	23	January 2006-December 2007	January 2008-April 2008

Chapter 2

University Quality and Labour Market Outcomes of Canadian Youth

2.1 Introduction

Investment in higher education is costly for a student in terms of direct financial resources (tuition and fees) and opportunity costs of forgone earnings. However, the long-term economic benefits of tertiary education are well documented. Based on expected returns, students decide whether to enter university and further, which university to attend. This paper explores the role that university choice has in the level of hourly wages during the initial transition from schooling to labour market following graduation. This is an important piece of information for high school graduates and their parents. I aim to answer the question of whether attending a more prestigious university has an independent effect on wage rate.

There is a vast literature on university quality that use European and American data. This begs the question of why we need a study with Canadian data in particular. The institutional structure of the EU and US countries is different from the Canadian education system in their admission policies, tuition fees and quality variation. European countries have a mainly public and tuition-free university system with admission policies that differ by country (standardized exam scores, high school grade point average). In the US, tuition fees vary a lot, the admission policy is the same accross the states (mainly based on SAT scores, essays and interviews). Meanwhile, in Canada there are 10 times less universities in number and the difference is quality is not as stark as among the US univeristies. Canadian universities are subsidized by the government and the tuition fees do not vary much. Given these differences, most of the findings of the American and European studies may not be generalized to the Canadian case. University quality is a topic not well analyzed for Canada, due in part due to the lack of Canadian data sets which identify the institutions explicitly.

This paper provides the first Canadian study that investigates and estimates the university ranking premium on starting wages. To our best knowledge, there is only one study conducted on Canadian data, Betts et al. (2007), which analyzes the relationship of several university characteristics with earnings in the late 1980's and early 1990's. This paper differs from Betts et al. (2007) in several ways. The matching methods employed coupled with the rich dataset tackle carefully the identification issues that arise in this setting. The approach used allows for nonlinearities in returns to university ranking. I use data from the older cohort of Youth in Transition Survey (YITS-B). An individual's university is directly observed in this micro data. The YITS, is one of the only two Canadian datasets[1] that has this information and it is also the most recent one among the two. Besides, the YITS contains detailed background information on the participants for the years 1998 to 2008. This feature of the data forms the basis of a credible identification strategy of the causal effect of university rankings on the starting wages of new graduates.

I use two different university rankings, a subjective and an objective ranking. The first is the "Best Overall Reputation Ranking" which is published by the *Maclean's magazine* and based on a survey the magazine conducts. I also construct a new ranking for Canadian universities that I call university quality ranking. This is another novelty of this study. Some papers use university characteristics like professor to student ratio, professor salaries, number of students, retention rate (Betts et al., 2007). Noticing a high correlation among the several university characteristics, other papers (Black and Smith, 2004, 2006) use factor analysis to combine them in one comprehensive index. Likewise, I use the principal component analysis to combine a set of different university traits, which signal different university attributes, into a single index. Based on this quality index I create a categorical rank variable.

Using matching methods I find that, controlling for a set of individual and family characteristics, there is a university ranking premium on the starting hourly wage rate of Cana-

[1]The other dataset that has information on the name of the universities that students atended is National Graduates Survey, Cohorts 1982, 1986, 1990. This is the dataset that Betts et al. (2007) use in their analysis.

dian Bachelor degree graduates. The university reputation premium of a top ranking university degree is on average 10% for women and 13.4% for men. When comparing the wages of the top 25% of the sample in the reputation ranking distribution to the bottom 25%, the returns are higher for both genders (15.2% for women and 29.9% for men). The ranking premium is higher for men than women and the results are robust through different specifications, samples and estimators. The results regarding the return to university quality are mixed. There is a 20.9% return for women when comparing two groups with a stark difference in university quality. For men I can only identify an estimate for the case when a high ranking university is defined as one with an above-median ranking. In that case, the return to university quality is 11.5% for men. The dose-response and the treatment effect functions indicate a nonlinear and positive relationship between university ranking and hourly wages. The wage returns per one rank increase are statistically significant only among the lowest ranking universities, which suggests that the employers notice even the small rank difference when negotiating wages with this group of graduates. Whereas in the case of the graduates from higher ranking universities (quality or reputation), employers seem to pool together this group and treat everyone the same without distinguishing the unit differences in ranking.

The paper is organized as follows. Having introduced the topic in this section, I review the existing literature in Section 2.2 and discuss the data in Sections 2.3 and 2.4, the methodology in Section 2.5, and the empirical results in Section 2.6. Section 2.7 concludes.

2.2 Literature Review

There is a vast literature that analyses the returns to education. Most of it is based on the Mincer (Mincer, 1958) earnings regression. As displayed in equation (2.1) below, the logarithmic wages are specified as a function of years of schooling, S_i, and a set of other individual characteristics, X.

$$log\,\omega_i = \alpha_0 + \alpha_1 S_i + X\beta + u_i \tag{2.1}$$

The coefficient α_1 is interpreted as the return to an additional year of schooling. Card (1999) reviews the contributions to this research area. He concentrates mainly on the papers that challenge two main implicit assumptions of the Mincer model: exogeneity of the years of schooling variable and the functional form. Firstly, the education variable in the above set up may capture other confounding effects of unobservable characteristics like the ability of the individual. Ability conveys important information about the behaviour process of the students. When not controlled for as it is an unobservable, it hides in the error term, $u_i = \gamma A_i + \epsilon_i$, where A_i is ability and ϵ_i is an independent error term. If there is not any way to control for A_i then $Cov(S_i, u_i) \neq 0$. Violation of this orthogonality assumption yields inconsistent estimates and unreliable hypothesis testing. This is because of two reasons: higher ability individuals pursue more education and also some of the observed wage premium these individuals get could be attributed to schooling when it is actually innate ability. Researchers have applied different methods to solve this problem. In the presence of rich data, some assume "selection on observable variables" and in that case the above equation takes the following form

$$log\,\omega_i = \alpha_0 + \alpha_1 S_i + \gamma A_i^* + X\beta + \epsilon_i$$

where A_i^* is a proxy measure of the latent ability (e.g.: high school grades, standardized test scores) and X includes other control variables (respondent's own background characteristics, experience and experience squared, family, friends and high school characteristics). In the information space of X and A_i^* the assumption $Cov(S_i, \epsilon_i) = 0$ holds and S_i is no longer endogenous in the empirical model. Due to a lack of information in the data, several other papers deal with selectivity by using instrumental variable techniques[2] to identify the returns to schooling from other confounding effects. Secondly, the assumption of the linear functional form of the Mincer equation is likely not to hold. The effect of education for the years 8, 12, 16 (coinciding to the completion of elementary school, high school and college or university) on the wage rate might be nonlinear - this is commonly known as the "sheepskin effect". Some non-linearities in those specific years of schooling might exist due to the fact that having completed a certain level of education and having obtained the diploma/certificate/degree documenting it, matters differently in the determination of a higher wage by the employee. What about the prestige of the institution that grants the

[2] A number of papers assume "selection on unobservables" and use proximity to college as an instrumental variable (IV) for years of education completed. Other instrumenting variables for S_i that are usually used are the education of the parents and the education of the partner/spouse. But those may be weak instruments for the university quality. Long (2008), instead, uses the average quality of the nearby colleges within a certain radius of the student as the instrument for the quality of the college that the student attends.

degree? Will that induce an additional increase in the wage rate beyond the education level attained? This is where the topic discussed in this paper fits in the Canadian literature of wage returns to education. Hence the above equation becomes:

$$log\, \omega_i = \alpha_0 + \alpha_1 Q_i^* + \gamma A_i^* + X\beta + \epsilon_i \tag{2.2}$$

where Q_i^* indicates the latent university quality variable. Our purpose is to examine the returns to the quality of the degree granting university, thus estimating parameter α_1. The research dedicated to analyzing the returns to university quality is extensive using US data, less so for European data and fairly new on Canadian data.[3] Black and Smith (2004, 2006) use US data and see the effect of the 4-year college quality on the hourly wage rate in 1989 and 1998. These two papers raise the issue of measurement error of the proxies used for the latent quality variable. They try to fix this issue by building a quality index using factor analysis and principal component analysis. Another way of dealing with measurement error is through instrumental variables. Black and Smith (2006) derive a GMM estimator, which they prefer best as opposed to factor analysis, because it makes direct use of the covariance matrix between the proxy variables. They find an average impact of 0.039 on the logarithmic hourly wage rate caused by one unit increase in the quality index. Black and Smith (2004) in a matching framework, where the quality variable is a binomial indicator

[3] Among the many relevant papers are Eliasson (2006); Chevalier and Conlon (2003); James et al. (1989); Brewer et al. (1998); Horstschraer (2011); Suhonen (2011); Heckman et al. (2003); Monks (2000); Dale and Krueger (2002).

of attending a high quality[4] four-year college, find an impact of 12-14% in the log hourly wage rate. Long (2008) criticizes this method reasoning that the amount of the observations not used (pertaining to the inter-quartile range) is big which reduces the sample size a lot and thus the efficiency in estimation. An other critique is related to the fact that the "*estimates refer to discrete moves from one group of universities to the other and do not allow the estimation of the effect of moving up the quality distribution within a group of colleges*" (Long, 2008, pg.594). Long (2010) looks into the trend of how the effect of years of education and four-year college quality changes over a period of 30 years (1970-2000) by using three different datasets that cover each of the three decades. He decomposes the analyses by gender and race and finds that the changes in the years of education and quality effects on a set of outcome variables are heterogenous among subgroups but mainly increasing through years for some of them. Black et al. (2005) also conduct a through-time analysis of the university quality on wage rates and find that it is quite stable during the time span 1987-1998 with men benefiting more than women (except in 1989). They also consider a few other labour market outcomes apart from the logarithmic hourly wage rate such as educational attainment, graduate school attendance, labour force participation, hours of work during the last year, marital status, number of children and spouse earnings.

Holmlund (2009) summarizes the studies on European data. In this paper the author contributes to the literature by analyzing a very rich Swedish dataset on individuals and university characteristics and by employing the quartile regression methods. She finds

[4] A four-year college here is defined as being "high quality" if it falls in the fourth quartile of the distribution of the quality index built by principal component analysis as opposed to falling in the first quartile.

that the returns to university quality are higher for the individuals who belong in the top quartiles of the income distribution.

To my best knowledge, the only paper that attempts to estimate alumni's wage returns to university traits in Canada is Betts et al. (2007). They use data from the National Graduates Survey and pool together three cross sections for the years 1982, 1986 and 1990. This dataset lacks a measure for the ability of the participants. In the absence of this important variable that could help in addressing the selection issue, the authors use a fixed effect model which *"to the extent that the most able students in a province always attended universities A and B over the eight-year period under study,[fixed effects] sweep average ability of the university's student body out of the wage equations"* (Betts et al., 2007, pg.10). The results are interpreted as *"something approaching a causal effect of [university] resources on student outcomes"* (Betts et al., 2007, pg.10). The outcomes of interest are labour force participation and annual earnings five years after graduation. Differently from Betts et al. (2007), I perform a cross-sectional analysis. There are a few strengths in the dataset that I use. The availability of a measure for ability as well as a wealth of information on individual and family characteristics, allow me to assume that selection into universities of different quality is based on these important observable variables. In this way, I am able to identify a possible causal effect of the university quality and reputation on hourly wage rates earned in the first job following graduation. Apart from being a very recent dataset, YITS is one of the two Canadian datasets so far that allow the identification of universities by name. Using this information, I match the individuals to external (not within YITS-B) data on the characteristics of the university that granted their Bachelor's degree.

2.3 The YITS Data

The main data set used in this paper is the Youth in Transition Survey Cohort B (YITS-B). Students of age 18 to 20 in December 1999 were surveyed every two years until April 2008, and each survey asks questions related to the past two years from the date of the interview[5]. In the first wave of data the students were 18-20 years old and this time corresponds to the age range in which most of them have graduated from high school and enrolled in a PSE institution. By the third wave the age range is 22-24. By this age, most of the students have graduated from at least a PSE program and are in the labour market. So, the outcome variable of interest is observed in the years 2003-2005.

Our sample focuses on the participants of YITS-B who have a Bachelor's degree or equivalent as of December 2003. This number does not include the individuals who attained university diploma or certificate below Bachelor's (undergraduate level). Because their wage structure is different from a regular BA degree, they are excluded from the sample. Also the sample excludes those individuals that have attained a professional degree, an Master's or a PHD degree. The reason is that their wages will be higher when compared to any BA graduate because of their post-graduate degree. Including these might confound

[5]Detailed information on the overall sample size, time of the interview, reference time and age of the participants can be found in Table 2.6 in the Appendix.

the university quality effect with that of a higher degree. This subsample contains 2,520 observations, 59% (or 1,485) of which are women and 41% (or 1,035) are men. Because the individuals must have an overall post-secondary status "graduate, non-continuer" as of December 2007, the subsample shrinks further to 2,026 (60% or 1220 women and 40% or 806 men). Of the observations deleted, 494 were those people that graduated from BA program but are continuing another BA program or a post graduate program. Since they are still students, they are not counted in the labour force. Hence, 20.4% of all participants in cycle 3 of the survey have completed and attained one BA degree from a Canadian university. Within this sample I am only interested in those students that have started a full-time or a part-time paid job as an employee; self-employed individuals, which are about 5% of the sample, are dropped from the analysis. This restriction and the unavailability of rankings for some universities reduces the sample even further to 672 women and 422 men.

YITS is a suitable dataset for the purpose of this paper because it contains a wealth of information. These include the overall high school grade point average that is a proxy for A_i^* in equation 2.2, field of study indicator variables, rural and provincial residence indicator of the respondent, respondent's number of dependent children, Canadian citizen indicator, full-time employee indicator, marital status indicator, respondent's father and mother post-secondary education indicator. Appendix A contains detailed description of the individual characteristic variables used in our specification.

The descriptive statistics for the individual characteristic variables can be found in Table 2.1. A lower fraction of women have parents with some post-secondary education than

Table 2.1: Descriptive statistics of individual characteristics

Variable	Mean	Standard Deviation	Obs. No.
Women			
Mother PSE	0.486	0.500	808
Father PSE	0.561	0.497	747
Rural Dummy	0.179	0.377	841
Number of dependent children	0.220	0.593	845
Citizen of Canada dummy	0.969	0.191	848
Married or Living with partner dummy	0.568	0.495	845
Separated/Divorced/Widow dummy	0.011	0.107	845
Men			
Mother PSE	0.573	0.495	565
Father PSE	0.627	0.484	536
Rural Dummy	0.121	0.326	572
Number of dependent children	0.133	0.428	590
Citizen of Canada dummy	0.938	0.241	592
Married or Living with partner dummy	0.486	0.500	590
Separated/Divorced/Widow dummy	0.007	0.107	590

men. The number of the respondent's dependent children reported is much higher for women than men, and they are more likely to have a non-single status.

2.4 University Ranking Data

I merge YITS-B with the university characteristics from external sources. The data on university[6] quality indicators are obtained from the publicly available data in the university ranking issue of the *Maclean's magazine* published on November 2002, and Canadian Association of University Teachers (CAUT) Almanac published in 2002.

In the analysis, I use two university rankings. The first is the *Maclean's magazine* Best

[6]The universities included in the sample are 45 and include: Universities of Toronto, McGill, British Columbia, Alberta, Queen's, McMaster, Dalhousie, Calgary, Western Ontario, Saskatchewan, Ottawa, Laval, Montreal, Sherbrooke, Manitoba, Simon Fraser, Victoria, Waterloo, Guelph, Memorial, New Brunswick, Carlton, Windsor, Regina, York, Concordia, Mount Allison, Acadia, Lethbridge, Wilfrid Laurier, Trent, St.Francis Xavier, Bishop's, Prince Edward Island, Winnipeg, Saint Mary's, Lakehead, Brock, Laurentian, Brandon, Ryerson, Mount Saint Vincent, Moncton, Cape Breton and Nipissing. The data for university characteristics for UQAM, UOIT, St.Thomas, UNBC were not sufficient to build the quality index and are dropped from the sample.

Overall Reputation Ranking which is constructed from the survey data that the *Maclean's magazine* collects every year. This survey asks high school counsellors, university officials, CEOs and corporate recruiters across Canada to separately rank the universities on three attributes: best quality, most innovative and leaders of tomorrow. Then, the "Best Overall Reputation Ranking" is calculated as an average of the three rankings. The second university ranking that I use, is an objective ranking that I construct from a set of university financial, physical and human capital inputs. Using the Principal Component Analysis, I combine different university characteristics in an index. The university characteristics include average high school grade of entering cohort, proportion of students who graduate, student-faculty ratio, number of full-time and part-time students, percent of classes taught by tenured faculty, percent of faculty with doctoral degrees, proportion of students who have won national awards in the past 5 years, proportion of faculty members who have won national awards in the past 5 years, size and number of peer-adjudicated grants per faculty member, university operating expenditures per full-time student, percentage of total operating expenditures devoted to student services and to scholarships and bursaries, number of volumes in university libraries per full-time student, the proportion of library budget allocated to updating the university's collection and to maintaining library services, the percentage of alumni who made gifts to the university in the past 5 years, tuition and fees. Detailed definitions are provided in Appendix A.

Principal Component Analysis (PCA) yields linear orthogonal combinations of the individual characteristics by assigning weights to each. The selected weights are optimal and maximize the ability of the index to capture the covariance between the university characteristics. PCA may create as many orthogonal combinations, known as components, as there are inputs, in this case the university characteristics. Starting with the first component, the extent of covariance accountability that the component captures decreases in the second one and so on. Within a component the variable contributing most to the covariance is weighted highest. I use the first principal component of the orthogonal transformation as our quality index. This is an efficient and optimal way of combining many university characteristics into one without worrying about the multicollinearity when these, otherwise, enter together in a regression equation.

The first principal component is based on the correlation matrix and not the covariance matrix of university characteristics. Unlike factor analysis, PCA is not scale invariant; the eigenvalues and eigenvectors of a covariance matrix differ from those of the associated correlation matrix. Usually, a PCA of a covariance matrix is meaningful only if the variables are expressed in the same units. The university variables have different units, thus I use correlation matrix instead. PCA is built based on two main assumptions that all the variables have a multivariate normal distribution and that the covariance matrix of the observations has all distinct and strictly positive eigenvalues. The first assumption can be tested by the multivariate normality test of Henze-Zirkler(1990). The university characteristics published in 2002 pass this test (p-value=0.285).

The *Maclean's magazine* Best Overall Reputation Ranking is a categorical variable ranging in value from 1 to 45, a higher number corresponding to a better ranking. The university quality ranking is constructed as a categorical variable, too, so that the results may be comparable among the two. The two rankings capture different attributes of universities. The reputation ranking is based mainly on the experience that the surveyed individuals have had with the different university graduates and subjective perceptions of the relative university quality. The quality ranking constructed by PCA is an objective measure of university quality because it is based on the human capital inputs (faculty and students) and allocation of financial resources which facilitate learning and promote research.

I use different forms of the treatment variable (university ranking). First, in one of the OLS specifications the ranking variables are used in their categorical form ranging in value from one (lowest rank) to 45 (highest rank). In this specification, the coefficient indicates the average returns to an increase in the rank of reputation (quality) by one rank. Second, allowing for nonlinearity in the returns to university ranking, I define the ranking variable as an indicator of value one if the university pertains to the fourth (or top) quartile of the university ranking distribution, and zero if it pertains to the first (or bottom) quartile. Referring to Figure 2.1 panel (a), the quality variable is a value of one for all the individuals within the area "Q.4" (i.e. quartile 4), and zero for the individuals in the area "Q.1". This definition of the treatment variable uses only the data in the tails of the distribution and thus 50% of the sample. In order to use all the data available, I define a second treatment variable (panel (b) in Figure 2.1) which determines the treatment as graduating from a university in the top 50% of the ranking distribution versus the bottom 50%. As it is common in this literature, I start the analysis with the ordinary least squares (OLS) estimates. Next, I use

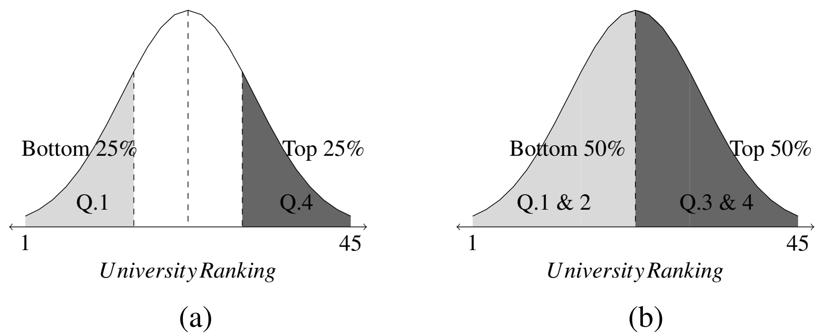


Figure 2.1: Visual representation of the ranking indicator variables

matching techniques to estimate returns to university quality: nearest neighbour matching and propensity score matching. There are several advantages in using matching methods relative to least squares (OLS) regression. First, least squares regression assumes the causal effect of the treatment is constant for each individual, while matching techniques estimate this effect for each individual i in the sample, and report the average of the individual effects. Second, unlike OLS, matching disposes of the assumption that the relationship between the treatment and the outcome of interest is linear. Third, the balancing property in OLS is assumed, whereas matching methods emphasize it and we can explicitly test for it (see Rosenbaum and Rubin, 1983)[7].

The descriptive statistics of university characteristics are shown in Table 2.7 in the Appendix. *Maclean's magazine* classifies the universities into three categories. The first category is the "Medical Doctoral" which includes those universities with a broad range of PhD programs and research as well as medical schools; the second category is the "Comprehensive" which includes those universities with a significant amount of research activity

[7] For a technical and detailed description on the matching techniques see Rosenbaum and Rubin (1985); Abadie and Imbens (2006); Cochran and Rubin (1973); Dehejia and Wahba (1999); Heckman et al. (1997, 1998b,a,c); Imbens (2000); Rosenbaum and Rubin (1983, 1985); Rubin (1974, 1980).

and a wide range of programs at the undergraduate and graduate levels including professional degrees; and the third category is the "Primarily Undergraduate" which includes those universities largely focused on undergraduate education with relatively few graduate programs. All university characteristics, are highest for the Medical/Doctoral universities and lowest for the Primarily Undergraduate universities, except for the Student Services, Alumni Support and Faculty to Student Ratio. By and large, Medical/Doctoral universities have the highest endowments and resources among the three categories.

2.5 Methodology

Let the outcome be $log\,\omega_i$ representing the log hourly wage of each individual at the first job after they graduate from a Bachelor's degree. The potential outcome, which is a different notion than the observed outcome, for each treatment state is

$$log\,\omega_i = \begin{cases} log\,\omega_{1i} & \text{if } H_i = 1 \\ log\,\omega_{0i} & \text{if } H_i = 0 \end{cases}$$

where H_i is a treatment dummy variable that takes a value of one if the individual graduated from a high rank university and zero otherwise. Our coefficient of interest is the average treatment effect on the treated (ATT) defined as

$$ATT = E\left[log\,\omega_{1i} \mid H_i = 1\right] - E\left[log\,\omega_{0i} \mid H_i = 1\right]$$

So, ATT is the average log hourly wage difference between those that graduated from a higher ranking university and the average log hourly wage that these same individuals would have had if they had graduated from a lower ranking university. The later is unobserved because we can not observe one same individual in both states, and thus we can not see both potential outcomes of an individual in the treatment and non-treatment case. $E\left[log\,\omega_{0i} \mid H_i = 1\right]$ is commonly known as the counterfactual. We can only estimate the counterfactual by $E\left[log\,\omega_{0i} \mid H_i = 0\right]$ and thus estimate ATT as the difference between the average outcome of the treated (higher ranking university graduates) and of those who were not treated (lower ranking university graduates). However, this is only possible at a cost. As shown in Angrist and Pischke (2008, pg.11) the equation below clearly displays this cost, the selection bias. Selection bias derives from the fact that students with certain attributes and background self-select into university education, and moreover self-select into the higher ranking universities. A solution to the selection problem is the random assignment of students to universities of different quality. This would make the two groups (treated and untreated) comparable and make possible the calculation of the counterfactual. Random assignment can be guaranteed when the data are experimental and the researcher has direct control on assigning the treatment randomly.

$$\begin{aligned}
\underbrace{E\left[log\,\omega_i \mid H_i = 1\right] - E\left[log\,\omega_i \mid H_i = 0\right]}_{\text{Observed Difference in Average Outcome}} \quad = \quad & \underbrace{E\left[log\,\omega_{1i} \mid H_i = 1\right] - E\left[log\,\omega_{0i} \mid H_i = 1\right]}_{ATT} \\
& + \\
& \underbrace{E\left[log\,\omega_{0i} \mid H_i = 1\right] - E\left[log\,\omega_{0i} \mid H_i = 0\right]}_{\text{Selection Bias}}
\end{aligned}$$

In the case of non-experimental data (e.g. survey data), researchers are able to assume that selection into universities is dependent on some characteristics which can be observed or measured like family background, own attributes, past academic performance, etc. This is commonly known as the selection-on-observables or conditional independence assumption (CIA). In notation: $log\, \omega_h \perp H \mid X$ for all $H \in \{0, 1\}$. What this says is that treatment is assigned "as if randomly" after conditioning on a sufficient set of variables based on which the individuals self-select or are selected by the universities. Under CIA, after conditioning on X, a matrix containing predetermined characteristics of individual i, we have

$$E\left[log\, \omega_{0i} \mid X_i,\; H_i = 1\right] = E\left[log\, \omega_{0i} \mid X_i,\; H_i = 0\right]$$

So, we can easily estimate the average treatment effect on the treated as

$$ATT = E\left[log\, \omega_{1i} \mid X_i,\; H_i = 1\right] - E\left[log\, \omega_{0i} \mid X_i,\; H_i = 0\right]$$

Nearest neighbour (NN) matching method calculates returns to education by finding for each treated individual at least one untreated individual that has the closest values of X_i as the treated individual and calculate the difference in their hourly earnings. After doing this for each treated individual, ATT is just the mean of all these differences as shown below:

$$\beta^{NNM} = \frac{1}{N_1} \sum_{i:H_i=1} (Y_i - \tilde{Y}_{0i})$$

where

$$\tilde{Y}_{0i} = \frac{1}{M(i)} \sum_{m \in M(i)} Y_{0m}$$

Abadie and Imbens (2002) show that in finite samples the matches may not be exact in their characteristics, and this creates a bias. Abadie and Imbens (2006) suggest a bias-correction adjustment based on a linear regression and show that it performs better than NNM without bias correction and ordinary least squares. Nearest Neighbour Matching estimator with Bias Correction (NNMBC) differs from the simple NNM method in the calculation of the counterfactual, $\tilde{Y}_{0i}$.

$$\tilde{Y}_{0i} = \frac{1}{M(i)} \sum_{m \in M(i)} (Y_{0m} + \hat{\mu}_0(x_i) - \hat{\mu}_0(x_m))$$

where the difference of $\hat{\mu}_0(x_i) = E\left[Y_{0m}|X = x_i\right]$ and $\hat{\mu}_0(x_m) = E\left[Y_{0m}|X = x_m\right]$ is added in the formula to adjust for any difference in observable variables, A_i^* and X_i, between the matched individuals.

One issue with NN matching is what the literature refers to as "curse of dimensionality". The more variables you include in X, the more you guarantee that CIA holds, however as the number of these variables increases the bigger the number of cells defined by the values of X get and each cell of the multivariate distribution of the treatment and X becomes less and less populated and some of these cells are even empty. When this happens, the calculation of the counterfactual is not possible.

Differently from NN matching, propensity score matching[8] (PSM), aiming to overcome the "curse of dimensionality" issue, calculates the counterfactual by matching the individuals on the probability of getting the treatment, known as the propensity score. In this way matching is done based on only one variable and it is less likely to have empty cells (shown by Rosenbaum and Rubin, 1983). For the PSM estimator, the CIA is represented as

$$log\, \omega_H \perp H \mid s(H, X) \text{ for all } H \in \{0, 1\}$$

where $s(H, X)$ is the propensity score and is defined as the conditional probability of receiving treatment H having certain pre-treatment characteristics X.

Building on Rosenbaum and Rubin (1983), Hirano and Imbens (2004) introduce the estimation of the propensity score in the case of a continuous treatment, named generalized propensity score (GPS). They then calculate the dose response function and treatment effect function based on the GPS. In the case where the treatment is a continuous and normally distributed variable, say Q^*, the CIA in notation is

$$log\, \omega_{q^*} \perp I(Q^* = q^*) \mid s(q^*, X) \text{ for all } q^* \in Q^*$$

where $I(.)$ is an indicator function and $s(q^*, X)$ is the GPS. The estimation of the dose-response function is done in two steps. First, the conditional expectation of the outcome as

[8] I use the *psmatch2* command in Stata of Leuven and Sianesi (2003).

a function of two scalar variables (university quality index and GPS level) is estimated,

$$\beta(q^*, s) = E\left[log\,\omega \mid Q^* = q^*, GPS = s\right]$$

Second, in order to estimate the dose-response function at each treatment level, the conditional expectation of the outcome estimated in step one is averaged over the score of GPS calculated at each particular level of the treatment, i.e.

$$\mu(q^*) = E\left[\beta(q^*, s(q^*, X))\right]$$

A version of this method was coded and provided for use as a package in STATA by Bia and Mattei (2008). For a detailed description of the functional form of $s(h, x)$, $\mu(q^*)$ and $\beta(q^*, s)$ see Hirano and Imbens (2004) and Bia and Mattei (2008).

2.6 Results

The participants of the YITS-B survey graduated from high school in 1999 and most of them applied for a university program. Hence, university rankings published in 1999 were particularly important for them when making the choice as to which university to attend. On the other hand, most of this cohort graduated from university in 2002 and 2003. So, the potential employers would look at the 2002 rankings in order to create an idea of the quality of a degree and differentiate among many applicants for a job application. Because

Table 2.2: Cross tabulation by high school GPA and university reputation

High School Grade(%)	Ranking quartiles				
	First Quartile	Second Quartile	Third Quartile	Fourth Quartile	Total
60-79	8.39	5.93	3.85	13.75	31.87
80-89	8.35	11.21	9.56	20.22	49.34
90-100	2.31	5.16	3.74	7.58	18.79
Total	19.01	22.31	17.14	41.45	100

Note: The numbers in each cell are the cell percentage determined by the university reputation ranking and ability. The total number of observations is 910.

the aim is to capture the university ranking premium in the initial transition to labor market, the 2002 ranking data is used.

The rich data availability in the YITS presents an opportunity to follow an identification strategy that may be viewed as a potential quasi-natural experiment. The data has a measure for the ability of the students, the high school grade point average. Students may select into universities of different reputation (quality) based on their ability. Table 2.2 shows the bivariate distribution of students conditional on university reputation and high school grades.[9]

As expected top quartile universities attract mainly middle and high ability students; the middle ranked groups of universities educate mostly middle-ability students; whereas, the lowest ranked group of universities attract middle and low ability students. Given that high ability is a potential source of selection, controlling for high school grades in our specifications helps identification. High school grade point average conveys important information about the behaviour process. This is because higher ability individuals go to better schools

[9]Note that the high school grades variable in our data is a categorical variable indicating a grade interval. Since the high school grades are self-reported, there is always the risk that they may be overstated. However, in YITS the students where asked to report a grade interval. This procedure diminishes significantly the risk of measurement error.

and some of the observed wage premium these individuals get could be attributed to university quality when it is in fact innate ability. Not conditioning on high school grades, leads to an estimated high quality premium. The lack of an ability measure in the specifications produces misleading results. Hence, the availability of a measure for ability, and other characteristic variables, help identify an unbiased estimate of quality premium.

The average hourly earnings of men and women who graduated from a high ranking university (top quartile) is $3.00 higher than those who graduated from a lowest ranking university (bottom quartile). In Figure 2.2, the empirical distributions of wages separately for the top and bottom ranking quartiles is plotted.

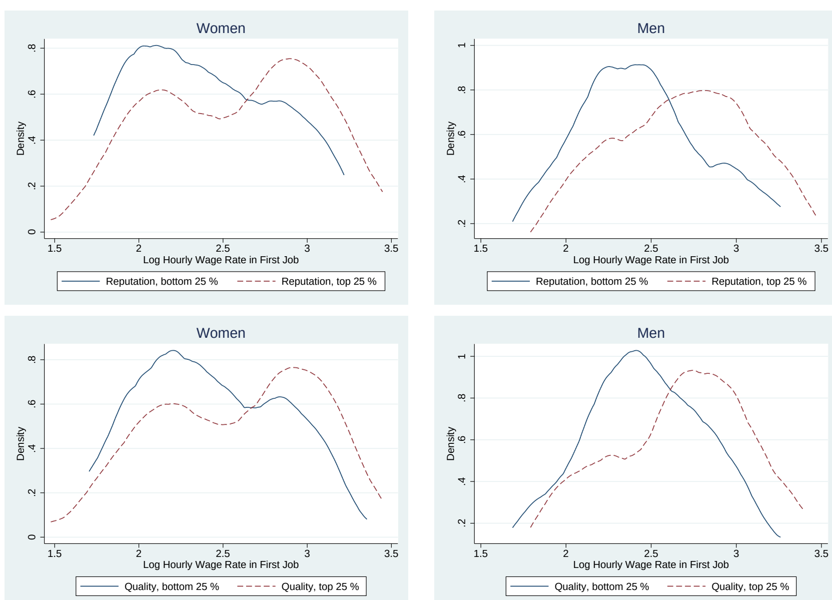


Figure 2.2: Empirical distribution of log hourly wages by gender, reputation and quality of university

In each of the panels, the wage density function is left skewed for the graduates of high ranking universities, and right skewed for those who graduated from lower ranking universities. This is true for reputation and quality rankings. For both genders there is a higher concentration of observations in higher wages for the graduates of selective universities. However, for those women that graduated from top ranking universities, their empirical distribution is bimodal with two clusters, one at the higher end and the other at the lower end of the wage distribution.

2.6.1 Quality Premium using Subjective *Maclean's* Ranking

The main sample includes full-time or part-time employees that graduated from a Bachelor's program and started working. Table 2.3 contains the results retrieved by using the nearest neighbour matching estimator with bias correction (BCNNM) for two definitions of the treatment variable (i) graduating from a university in the top 25% versus bottom 25% of the reputation ranking distribution, and (ii) graduating from a university in the top 50% versus bottom 50% of the reputation ranking distribution. In all these specifications I condition on high school grades, Bachelor degree field of study and other individual characteristics.

Table 2.3: Returns to university reputation ranking, BCNNM

Y=log(hourly wage)	Top vs. Bottom 25%	Top vs. Bottom 50%
Women	0.152**	0.103**
	(0.061)	(0.045)
	322	620
Men	0.299***	0.134***
	(0.084)	(0.051)
	214	400

Standard errors in parenthesis. ***Significance at 1%, **Significance at 5%, *Significance at 10%. The third number in each cell is sample size. The sample includes only the individuals that graduated from a Bachelor's degree and started a job full-time or part-time. BCNNM stands for Bias Correction Nearest Neighbour Matching.

The estimates show that female employees, who graduate from a university in the top quartile of the reputation ranking distribution earn a 15.2% increase in the hourly wage when compared to those females with the same measured ability and socio-economic characteristics that graduated from the bottom quartile university. Men earn almost twice as high premium (29.9%) as women. The reputation premium drops in magnitude when the full sample of Bachelor degree graduates is used, that is when the observations in the inter-quartile range of the ranking distribution are included. On average, reputation ranking premium for females is 10.3% and for males is 13.4%.

2.6.2 Quality Premium using Objective Rankings

In this subsection the above analysis is repeated but with a different ranking variable, the objective quality ranking constructed by PCA. This ranking is a proxy for university quality; it is a measure that relies on the amount of inputs of universities which translate into facilities and opportunities for their students. The inputs include several indicators of student body composition, faculty qualification and achieved grants, and lastly financial resource allocation. This would distinguish them in the labor market by the knowledge and skills that they posses rather then by the reputation of their degree.

Table 2.4 in the Appendix displays the return to university quality estimated by the nearest neighbour matching approach with bias correction. We can see that for women there are statistically significant returns to having a Bachelor's degree from a university that belongs to the top quartile of the university quality distribution. Women with higher quality

Table 2.4: Return to university quality rankings, BCNNM

Y=log(hourly wage)	Top vs. Bottom 25%	Top vs. Bottom 50%
Women	0.209***	0.009
	(0.057)	(0.047)
	332	620
Men	-0.031	0.115**
	(0.081)	(0.053)
	193	400

Standard errors in parenthesis. ***Significance at 1%, **Significance at 5%, *Significance at 10%. The third number in each cell is sample size. The sample includes only the individuals that graduated from a Bachelor's degree and started a job full-time or part-time. BCNNM stands for Bias Correction Nearest Neighbour Matching.

education earn 20.9% more than those who received a lower quality education. Notice that in this case the difference in the treatment received from the treatment and control group is highest, thus the return is expected to be high. When the treatment variable aims to compare the top 50% with the bottom 50% of the quality distribution, the return for women is not statistically significant and small in magnitude. The reason is that the inclusion of the inter-quartile range of the university quality distribution dilutes the difference in wages.

There seems to be no returns to attending a top quartile quality university versus a bottom quartile quality university for men. The sample size is much smaller and the standard error of the estimate is high. However, when comparing the top half with the bottom half of the university quality distribution, the return of attending a high quality university is 11.5% for men. This results may be related to the extend that the index distinguishes among the universities that are in the margins of the cut-off in the ranking distribution tails. Some of the universities that reputation ranking includes as top 25% ranking universities, the quality ranking misses. This may be seen better from Figure 2.3, where the two ranking variables (higher values are higher ranks) are plotted against each other. Each dot in the scatter plot represents a university and the line is a 45 degree line. The universities that are positioned on the 45 degree line, are assigned the same ranking by both measures. Those lying above the line are assigned a higher quality ranking, and those below the line are assigned a higher reputation ranking.

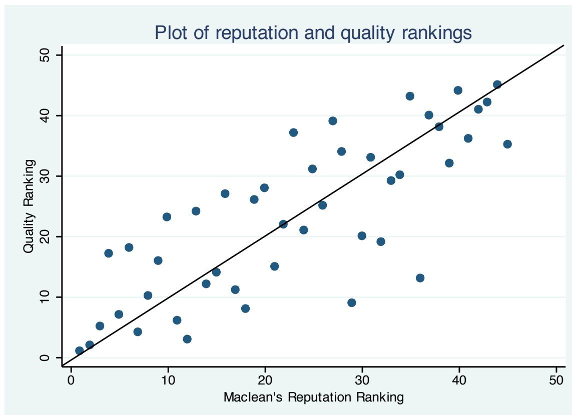


Figure 2.3: A scatter plot of reputation and quality ranking for each university

Notice that even though the two different rankings are highly correlated (Pearson rank correlation of 0.78), there is a high dispersion for the middle ranked universities. This indicates that while the rank of the very high quality and very low quality university are pinned down easily by both rankings (closer to the 45 degree line), it is harder to distinguish the among the middle ranked universities (higher dispersion away from the 45 degree line).

2.6.3 Rankings as Categorical Treatment Variable

The specifications presented in this section pool together men and women, and condition on a gender indicator variable. Figures 2.4 and 2.5 show the dose-response function and the treatment effect function for the pooled sample of females and males. The dose-response function indicates a non-linear, positive relationship between the university ranking and wage rates.

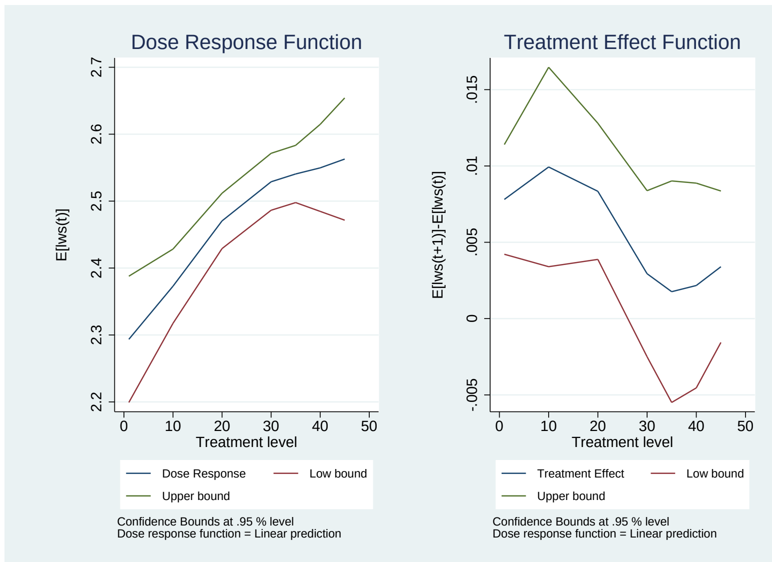


Figure 2.4: Dose-response and treatment effect functions - Reputation Ranking

The graph is steeper for the lower ranks and becomes flatter as the ranking levels increase. This may be seen more easily from the treatment effect function which shows a marginal increase per one rank improvement to be significantly different from zero for treatment levels between 1-25. The marginal increase in wages is statistically zero for a unit increase in ranking within the higher ranking universities (between 26-45). These results suggest that the graduates from lower ranking universities are those who benefit even for very small changes in ranking. The employers notice even one rank difference in negotiating wages. Whereas in the case of graduates from higher ranking universities (quality

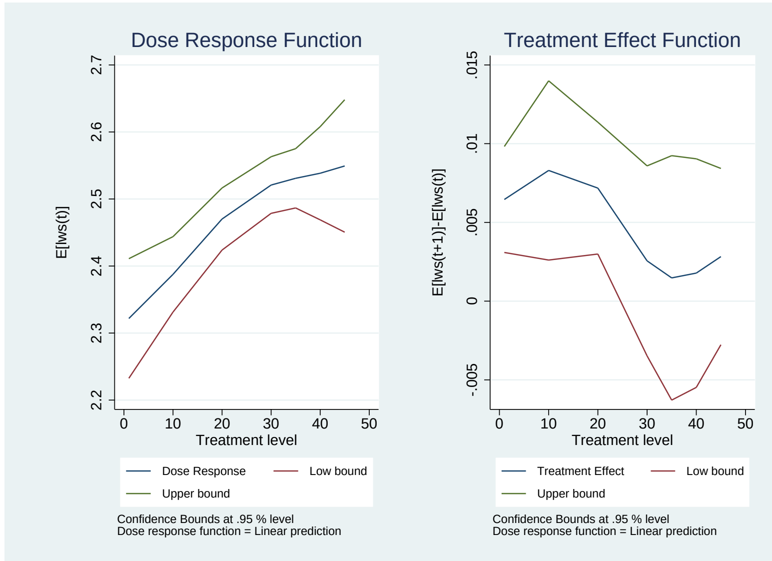


Figure 2.5: Dose-response and treatment effect functions - Quality Ranking

or reputation), they seem to be pooled together in one group and the employers do not distinguish the small differences in university ranking for these graduates. This result is consistent with the findings in Lang and Siniver (2011).

2.6.4 Sensitivity Analysis

The previous sections discussed the returns estimated by the Nearest Neighbour matching (NNM) method with bias correction. The reason why this estimator is preferred is twofold. First, NNM methods make use of few but very close matches when compared to Propensity Score matching (PSM) which may provide a few more in number but lower quality matches. This property makes NNM a relatively more efficient estimator. Second, Abadie and Imbens (2002) show that in finite samples the matches may not be exact in their characteristics, and this creates a bias. Abadie and Imbens (2006) suggest a bias-correction adjustment based on a linear regression and show that it performs better than NNM without bias correction and ordinary least squares.

Tables 2.8 and 2.9 in the appendix replicate our results using different estimators. The covariates in each specification include the measure for ability (Overall high school GPA), rural versus urban residence dummy while in high school, rural versus urban residence dummy while full-time employed, number of the respondent's dependent children, citizen status dummy, marital status, province of residence dummies, parental education dummies and undergraduate degree field of study indicators. Each of the cells in these two tables contains the parameter estimate and the standard error (see details in the table footnote). Referring to column (2) in Table 2.8 , it can be seen that the results presented in the previous two sub-sections are independent of the methodology used. Even though the estimated magnitude varies a little, the statistical significance is almost the same always indicating significant returns to university reputation ranking. Note that NNM with bias-correction estimates are more efficient than the Propensity Score Matching estimates.

In Table 2.8 different sample exclusions are considered (column (1) to (4)). One could argue that the best students with a deep interest in the subjects most likely will continue post-graduate education and attend a master's degree or choose to attain a professional degree. For this reason consider the differences in BCNNM estimates between columns (1) - (3) and (2) - (4). As the students that have finished a professional degree or a master's degree are included in the sample, the reputation premium estimates decrease for women but increase for men. This decrease in magnitude for women may be attributed to the fact that high ability graduates from low-ranking universities will pursue a graduate degree that will complement their skills. In this way the difference between the Bachelor degree graduates with higher versus lower reputation diminishes. For men, the contrary seems to happen. In this case, high ability graduates from high-ranking universities will pursue a graduate degree increasing further the premium of university ranking. However, this speculation should not be interpreted as selection to graduate school. Table 2.5 shows that the probability of attending a graduate program, conditional on ability and a set of demographic characteristics, is not affected by the reputation (or quality) ranking of the university where undergraduate studies were completed. Thus, this results indicate that there is no selection to graduate school caused by the reputation or quality ranking of the university. One other source of selection might be self-employment. In YITS-B only 5% of the students choose to be self employed. This is a small number of observations that most likely will not effect the results in this paper.

Table 2.5: University rankings and probability to attend a professional or graduate degree

	Top vs. Bottom 25%		Top vs. Bottom 50%	
Treatment Variable:	Reputation Ranking	Quality Ranking	Reputation Ranking	Quality Ranking
Women	0.037	0.086	0.002	0.045
	(0.194)	(0.213)	(0.148)	(0.155)
Men	-0.268	0.258	-0.106	-0.055
	(0.255)	(0.279)	(0.193)	(0.201)

Note: Each cell represents a *probit* regression and reports only the marginal effect estimate of the treatment variable, conditional on the high school grades and the demographic characteristics listed in table 2.1. Standard errors are in parenthesis. ***Significance at 1%, **Significance at 5%, *Significance at 10%.

Lastly, even though a measure for ability goes a long way in identifying a causal effect of university rankings on wages, I acknowledge that there may be other unobservable traits of individuals that could determines their decision-making. One such variable is motivation. In the YITS-B questionnaire there is a question about the aspirations of students. More specifically, they are asked the highest level of education they "would like to get", differentiating from the question "as things stand now, what is the highest level of education you plan to get?". A further question asks them if motivation is standing in their way of attaining the level of education they would like to achieve. I use these two questions to create a variable that would capture the motivation and preferences of the students. The results have negligible changes in magnitude after controlling for this variable.

2.7 Conclusion

This paper estimates the wage returns of university reputation and quality rankings for Canadian youth. Two university rankings are used. One is the *Maclean's magazine* Best Overall Reputation Ranking. I also build a new university ranking based on a quality index that I construct as the principal component of the Principal Component Analysis of several university characteristics. University characteristics data are retrieved from the *Maclean's magazine* November 2002 issue and CAUT Almanac 2002 issue. The analysis is split by gender. Our main data source is Youth in Transition Survey and the outcome variable of interest is the starting hourly wage in the first job after graduating from university. Several main findings emerge from the analysis in this paper. Firstly, I observe that the lack of an ability measure, individual and parental characteristics in the specifications would produce misleading results. The findings indicate that university quality matters a lot for both genders when I do not control for high school grade point average (GPA), which in turn conveys important information about the behaviour process. This is because of two reasons: higher ability individuals go to better schools and some of the observed wage premium these individuals get could be attributed to university quality when it is actually innate ability. Hence, the availability of a measure for ability helps identify an unbiased estimate for the returns to education quality.

Matching methods are employed and a sensitivity analysis through different estimators and sample exclusions is provided. The findings indicate that university reputation premium to graduating from a top ranking university is 10% for women and 13.4% for men. When comparing the wages of the top 25% of the sample in the reputation ranking distribution to the bottom 25%, thus excluding the middle-ranking observations, the returns are higher for both genders: 15.2% for women and 29.9% for men. The ranking premiums are higher for men than women and the results are robust through different specifications, samples and estimators. The dose-response and the treatment effect functions indicate a nonlinear and positive relationship between university ranking and hourly wages. The wage returns per one rank-increase are statistically significant only among the graduates of lower ranking universities, which suggests that the employers notice even the small rank difference when negotiating wages with this group of graduates. Whereas in the case of the graduates from higher ranking universities (quality or reputation), employers seem to pool together this group and treat everyone the same without distinguishing the unit differences in ranking. This results coincides with other studies in the literature.

The results obtained from the analyses in this paper may be of practical use to the students and their parents when making a life-changing decision regarding their post-secondary education track. Given the significant returns to university quality, this study also raises awareness of the potential consequences of student-sorting by university quality. From a policy perspective point of view, this is of particular importance when considering the possible long-run effects on the labour market and the income distribution.

The analysis in the present paper may be extended further in several aspects. First, I plan to see whether the university ranking affects other outcomes like yearly earnings and other benefits in the job, satisfaction in the job, probability to drop out of university and probability to graduate. A last extension to consider is building a better university quality index that would take into account that universities may rank differently based on field of study. In this paper I can only look at the short term effects of university quality and reputation. It would be very interesting to study the long-term effects and, in particular, to test if reputation or quality of the university would take individuals to different wage profiles.

Appendix A: Variable Definitions

Definitions of university characteristics

Dependent Variable

Log hourly wage: Logarithmic hourly wage paid when first hired (first job) after graduating university.

Personal Characteristics

Overall high school GPA: The overall high school grade point average (GPA). This variable is reported in intervals of 10 percentage points. Hence, it is a categorical variable.

Bachelor Degree Field of Study: A set of 11 dummy variables indicating the undergraduate field of study which include: (1) Education; (2) Visual and Performing Arts, and Communications Technologies; (3) Humanities; (4) Social and Behavioural Sciences, and Law; (5) Business, Management and Public Administration; (6) Physical and Life Sciences, and Technologies; (7) Mathematics, Computer and Information Sciences; (8) Architecture, Engineering and Related Technologies ; (9) Agriculture, Natural Resources and Conservation; (10) Health, Parks, Recreation and Fitness; (11)Personal, Protective and Transportation Services.

Rural Dummy: Indicator of rural or urban geography of the most recent residence of the survey participant. This is derived based on the Statistical Area Classification (SATYPE) 2001 Census geography.

Number of Children: Number of the dependent children of the respondent.

Citizen Dummy: Indicator variable takes the value 1 if the respondent is a Canadian citizen and 0 otherwise.

Full-time Dummy: Indicator variable takes the value 1 if the respondent is working full-time in his first job after graduating from university and 0 otherwise.

Professional degree and Master's Dummy: Indicator variable takes the value 1 if the respondent has graduated from a professional degree or Master's before starting their first job after full-time schooling and 0 otherwise.

Marital Status: A dummy variable is generated for each "married and/or living with partner" and "separated, divorced or widowed". The omitted category is "single".

Residential Province Dummies: A dummy variable is generated as an indicator variable for each of the Canadian regions: Atlantic, BC, Manitoba and Saskatchewan, Alberta, Quebec and the other provinces. The omitted category is Ontario.

Parental Variables

Father PSE: A dummy variable indicating that father has some post-secondary education: "College", "University and Professional Degrees", "Graduate Degree". Omitted category is "high school or less than high school education".

Mother PSE: A dummy variable indicating that mother has some post-secondary education: "College", "University and Professional Degrees", "Graduate Degree". Omitted category is "high school or less than high school education".

Quality Measures

Quality Ranking 2002: First Principal Component of the Principal Component Analysis (PCA) of the university characteristics.

Reputation Ranking: *Maclean's* editors solicited the opinion of 5,467 high-school guidance counsellors, university officials, CEOs and corporate recruiters across Canada. The reputation survey of *Maclean's* is both regional and national in character, dividing the country into the following areas: the Atlantic provinces, Quebec, Ontario, and the four Western provinces. All respondents completed a national survey; university officials and guidance counsellors also completed regional surveys. The respondents rank the universities as the Highest Quality, as the Most Innovative, and as Leaders of Tomorrow. The *Maclean's magazine* calculates a Best Overall Reputation Ranking by weighting equally the rankings for the three attributes.

Proportion who graduate: Percentage of full-time second-year undergraduates who completed their degree within one year of the expected graduation date.

Classes Taught by Tenured Faculty: The percentage of first-year classes taught by tenured or tenure-track professors.

Faculty with PhDs: Percentage of full-time faculty with a PhD degree.

Average Entering Grade: The average final-year grades of freshman students entering from high school or Quebec's CEGEP system.

Student Awards: The five-year tally of the number of students, per 1,000, who have won national awards.

Faculty Awards: The five-year tally of the number of full-time professors, per 1,000, who have won national awards.

Faculty Social Sciences and Humanities Grants(SSHR): The average size and number of peer-adjudicated research grants from both the Social Sciences and Humanities Research Council and the Canada Council. The size of grants is listed per eligible full-time faculty member; the number of grants is per 100 eligible full-time faculty members. The ranking reflects a weighted average of the two.

Medical Science Grants(MedSci): The average size and number of peer-adjudicated research grants from both the Natural Sciences and Engineering Research Council and the Medical Research Council. The size of grants is listed per eligible full-time faculty member; the number of grants is per 100 eligible average of the two.

Operating Budget: These figures show the size of operating expenditures per weighted full-time-equivalent student.

Student Services: Percentage of total operating expenditures devoted to student services.

Scholarships & Bursaries: Percentage of total operating expenditures devoted to scholarships and bursaries.

Library Holdings per Student: These figures show the number of volumes in all campus libraries, divided by the number of full-time-equivalent students.

Library Acquisitions: The proportion of the library budget allocated to updating the university's collection.

Library Expenses: The percentage of the university budget devoted to maintaining library services.

Alumni Support: The percentage of alumni who made gifts to the university over a five-year period.

Student Faculty Ratio: The ratio of the number of full-time tenured faculty members to the number of students enrolled in an university. These data are collected from the yearly publication of CAUT Almanac.

Number of Full-time Students: Number of full-time students in a university.

Number of Part-time Students: Number of part-time students in a university.

Tuition: Tuition fee for Bachelor of Arts programs.

Compulsory and Ancillary Fees: Other fees that are paid additional to tuition fees for Bachelor of Arts programs.

Appendix B

Table 2.6: Timing of cycles for YITS - B and overall sample size

	Obs	Participants Age	Refence Time Period	Time of the Interview
Cycle 1	22,378	18-20	Jan1998-Dec1999	Jan2000-Apr2000
Cycle 2	18,779	20-22	Jan2000-Dec2001	Jan2002-Apr2002
Cycle 3	14,817	22-24	Jan2002-Dec2003	Jan2004-Apr2004
Cycle 4	12,435	24-26	Jan2004-Dec2005	Jan2006-Apr2006
Cycle 5	9,946	26-28	Jan2006-Dec2007	Jan2008-Apr2008

Table 2.7: Descriptive statistics for university characteristics

Variable 2002	Category	Mean	Std. Dev.	Min	Max	Obs
Average Entering Grade	1	84.867	1.959	82.000	89.000	15
	2	81.000	3.376	75.000	86.000	11
	3	79.619	2.636	76.000	85.000	19
Faculty with PhDs	1	94.607	2.840	88.800	98.400	15
	2	90.836	5.682	76.700	97.200	11
	3	81.971	13.209	38.200	95.300	19
Classes Taught by Tenured Faculty	1	56.620	9.546	38.700	72.600	15
	2	56.427	13.345	38.200	81.000	11
	3	63.557	11.282	39.100	85.700	19
Proportion who Graduate	1	86.853	5.710	72.600	92.900	15
	2	78.882	5.356	69.900	88.100	11
	3	76.710	9.213	54.000	92.300	19
Tuition	1	3338.333	1145.014	1663.000	4860.000	15
	2	3503.273	859.207	1668.000	4265.000	11
	3	4023.286	1027.281	1668.000	6584.000	19
Compulsory and Ancillary Fees	1	499.000	256.840	222.000	1143.000	15
	2	471.364	200.355	203.000	807.000	11
	3	427.429	236.353	65.000	892.000	19
Student Awards	1	5.733	1.852	2.500	9.500	15
	2	3.982	1.658	1.400	6.500	11
	3	1.795	1.247	0.200	4.300	19
SSHR Size	1	8420.200	3583.352	2780.000	14353.000	15
	2	4622.273	1973.140	1947.000	8502.000	11
	3	2297.238	1331.677	235.000	4967.000	19
SSHR Number	1	26.437	11.776	10.480	47.730	15
	2	16.431	8.347	7.220	33.950	11
	3	9.696	5.138	1.500	21.850	19
MedSci Size	1	66920.470	22555.350	24486.000	106137.000	15
	2	43145.450	18467.220	22248.000	80531.000	11
	3	13147.190	8840.186	0.000	34930.000	19
MedSci Number	1	119.352	34.513	62.650	194.000	15
	2	109.886	34.551	53.170	165.450	11
	3	51.153	25.624	0.000	92.540	19
Scholarships and Bursaries	1	9.429	2.596	4.750	13.690	15
	2	6.756	2.279	3.990	10.790	11
	3	4.893	2.345	1.450	9.180	19
Student Services	1	4.713	0.952	3.440	6.890	15
	2	5.119	1.101	3.880	7.440	11
	3	6.350	1.876	4.190	10.530	19
Library Acquisitions	1	45.818	4.774	37.960	51.180	15
	2	39.887	6.258	29.270	50.450	11
	3	36.438	5.563	28.290	49.340	19
Library Expenses	1	6.467	1.257	4.670	9.370	15
	2	6.510	0.656	5.350	7.470	11
	3	5.518	0.981	3.820	7.460	19
Library Holdings per Students	1	228.600	63.201	145.000	349.000	15
	2	213.091	57.146	133.000	312.000	11
	3	224.667	74.374	74.000	364.000	19
Faculty Awards	1	5.960	2.782	2.500	10.700	15
	2	3.455	2.490	0.500	8.200	11
	3	1.843	2.251	0.000	8.800	19
Operating Budget	1	8398.400	983.984	6981.000	10736.000	15
	2	7857.273	666.390	6741.000	8767.000	11
	3	7380.286	1186.084	4827.000	9533.000	19
Full Time Students	1	21698.330	9016.881	9003.000	44126.000	15
	2	14131.640	6265.645	7149.000	30056.000	11
	3	4184.667	2416.590	1887.000	11163.000	19
Part Time Students	1	7186.200	4568.307	2211.000	16728.000	15
	2	5330.455	3503.734	1512.000	12546.000	11
	3	1942.000	2506.355	281.000	12047.000	19
Alumni Support	1	16.653	4.339	9.800	25.400	15
	2	13.200	4.144	4.600	20.400	11
	3	18.350	7.494	3.200	32.500	19
Faculty Student Ratio	1	0.1733	0.029	0.100	0.220	15
	2	0.225	0.034	0.160	0.280	11
	3	0.202	0.048	0.130	0.270	19

Table 2.8: Starting wage returns to university reputation rankings

Y=log(w/h)	(1)		(2)		(3)		(4)	
	F	M	F	M	F	M	F	M
OLS								
Cont.Var	0.005*	0.006*	0.005***	0.006***	0.004**	0.007***	0.004***	0.007***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
4v1	0.195***	0.182**	0.171***	0.191***	0.153***	0.213***	0.131***	0.198***
	(0.062)	(0.073)	(0.053)	(0.063)	(0.057)	(0.066)	(0.048)	(0.057)
43v21	0.102**	0.141***	0.070*	0.138***	0.088**	0.163***	0.069*	0.150***
	(0.045)	(0.052)	(0.039)	(0.044)	(0.042)	(0.048)	(0.036)	(0.041)
NNM								
4v1	0.265***	0.234***	0.215***	0.227***	0.243***	0.231***	0.215***	0.198***
	(0.069)	(0.078)	(0.057)	(0.076)	(0.065)	(0.074)	(0.054)	(0.068)
43v21	0.140***	0.203***	0.109**	.183***	0.143***	0.199***	0.124***	0.169***
	(0.049)	(0.058)	(0.045)	(0.053)	(0.046)	(0.054)	(0.042)	(0.049)
BCNNM								
4v1	0.239***	0.074	0.152**	0.299***	0.196***	0.207**	0.138**	0.355***
	(0.071)	(0.087)	(0.061)	(0.084)	(0.068)	(0.087)	(0.056)	(0.068)
43v21	0.139***	0.127**	0.103**	0.134***	0.134***	0.130**	0.091**	0.150***
	(0.050)	(0.056)	(0.045)	(0.051)	(0.047)	(0.055)	(0.043)	(0.048)
PSM								
4v1	0.264***	0.225**	0.157**	0.254***	0.225***	0.21***	0.151***	0.263***
	(0.079)	(0.119)	(0.068)	(0.102)	(0.070)	(0.069)	(0.062)	(0.088)
43v21	0.138***	0.109*	0.090**	0.114**	0.123***	0.164***	0.089**	0.138***
	(0.057)	(0.073)	(0.048)	(0.06)	(0.053)	(0.052)	(0.043)	(0.054)
OBS								
4v1	250	187	322	214	283	202	425	273
Full	417	310	620	400	514	359	748	466

Note: Standard errors in parenthesis. ***Significance at 1%, **Significance at 5%, *Significance at 10%.
Estimators used are Ordinary Least Squares (OLS), Nearest Neighbour Matching (NNM), Nearest Neighbour Matching with Bias Correction (NNM BCE) and Propensity Score Matching (PSM).
Sub-sample (1) includes the Bachelor degree graduates with full-time employment post-graduation. Sub-sample (2) includes the Bachelor's degree graduates with a full or part-time employment post-graduation. In all the specifications we add a full-time employee indicator variable. Sub-sample (3) includes the Bachelor, professional or Master's degree graduates with full-time employment post-graduation. In all the specifications we add a professional and Master's degree indicator variable. Sub-sample (4) includes the Bachelor, professional or Master's degree graduates with full or part-time employment post-graduation. In all the specifications we add a full-time indicator variable and a professional degree and Master's indicator variable.

Table 2.9: Starting wage returns to university quality rankings

Y=log(w/h)	(1)		(2)		(3)		(4)	
	F	M	F	M	F	M	F	M
OLS								
Cont.Var	0.004**	0.004	0.005***	0.004*	0.003*	0.005**	0.004***	0.005**
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
4v1	0.147*	0.146*	0.125**	0.145**	0.125**	0 .175**	0.120**	0.161**
	(0.067)	(0.083)	(0.058)	(0.072)	(0.062	(0.075)	(0.054)	(0.065)
43v21	0.159***	0.134**	0.117***	0.141***	0.133***	0.145***	0.109***	0.142***
	(0.047)	(0.055)	(0.041)	(0.047)	(0.043	(0.051)	(0.037)	(0.043)
NNM								
4v1	0.123*	0.129	0.139**	0.109*	0.122**	0.144*	0.147***	0.119*
	(0.067)	(0.082)	(0.058)	(0.077)	(0.062)	(0.079)	(0.054)	(0.073)
43v21	0.141***	0.157***	0.108**	0.118**	0.137***	0.145***	0.112***	0.109**
	(0.050)	(0.059)	(0.045	(0.054)	(0.046)	(0.055)	(0.042)	(0.051)
NNM BCE								
4v1	0.179***	-0.117	0.209***	-0.031	0.08	0.033	0.107*	0.093
	(0.068)	(0.094)	(0.057)	(0.081)	(0.066)	(0.102)	(0.056)	(0.087)
43v21	0.097*	0.08	0.009	0.115**	0.129***	0.115**	0.035	0.133***
	(0.054)	(0.059)	(0.047)	(0.053)	(0.048)	(0.056)	(0.044)	(0.049)
PSM								
4v1	0.269***	0.179*	0.207**	0.041	0.09	0.054	0.133*	0.124
	(0.091)	(0.119)	(0.090)	(0.125)	(0.072)	(0.118)	(0.102)	(0.120)
43v21	0.181***	0.096*	0.111**	0.092*	0.195***	0.098*	0.109**	0.110***
	(0.059)	(0.069)	(0.053)	(0.063)	(0.071)	(0.061)	(0.560)	(0.054)
OBS								
4v1 Sample	227	145	332	193	281	173	407	228
Full Sample	456	327	620	400	514	359	748	466

Note: Standard errors in parenthesis. ***Significance at 1%, **Significance at 5%, *Significance at 10%.

Estimators used are Ordinary Least Squares (OLS), Nearest Neighbour Matching (NNM), Nearest Neighbour Matching with Bias Correction (NNM BCE) and Propensity Score Matching (PSM).

Sub-sample (1) includes the Bachelor degree graduates with full-time employment post-graduation. Sub-sample (2) includes the Bachelor's degree graduates with a full or part-time employment post-graduation. In all the specifications we add a full-time employee indicator variable. Sub-sample (3) includes the Bachelor, professional or Master's degree graduates with full-time employment post-graduation. In all the specifications we add a professional and Master's degree indicator variable. Sub-sample (4) includes the Bachelor, professional or Master's degree graduates with full or part-time employment post-graduation. In all the specifications we add a full-time indicator variable and a professional degree and Master's indicator variable.

Chapter 3

Tertiary Education Choices as Response to Unexpected Information

3.1 Introduction

Students choose to invest on post-secondary education (PSE) because they expect more education to be associated with higher future earnings (Card, 1999), as well as better non-pecuniary benefits on the job (Oreopoulos and Salvanes, 2009). While the long-term expectations on the labour market prospects influence student's decision to access PSE, some of them end up dropping out or graduating from a different major and/or different institution from the one they started. In fact, 12 percent of American students switch major after their first academic year at university (Zafar, 2011), 18 percent of students who attended college in the US in 1993 switched major (Arcidiacono, 2004), 40 percent of the students from low-income families drop out of college (Stinebrickner and Stinebrickner, 2012). In this paper I use data from the Youth in Transition Survey (YITS) to study this phenomenon. In the YITS, more than 18 percent of the students who started a post-secondary education (PSE) program in 2002-2003, report to have changed program after their first year in PSE, and about 10 percent report to have switched major within their institution.

Table 3.1: Percent of students switching major by first major

	Switched Major			Row % Total
First Major	EAH	STEM	SSBA	(Obs. #)
EAH	49.02	18.63	32.35	100
	(50)	(19)	(33)	(102)
STEM	34.11	47.19	18.6	100
	(44)	(61)	(24)	(129)
SSBA	29.31	34.48	36.21	100
	(17)	(20)	(21)	(58)
Total	(111)	(100)	(78)	(299)

Note: The numbers are percentages of students that have switched and reflect the transition from one major grouping to the other. Observation numbers are in parenthesis. EAH grouping includes majors in Education, Arts and Humanities. STEM grouping includes Sciences (Agricultural Biology, Health, Physical Sciences), Technology and Trades, Math and Engineering. SSBA grouping includes majors in Social Sciences and Business Administration.

The statistics shown in Table 3.1 indicate that among the students who report to have switched major after first year PSE, only about half (diagonal entries) switch to a major that is within the same grouping and not too different from the first. For example, few students (18.63 percent) from the the Education, Arts, Humanities (EAH) majors switch to STEM (Sciences, Technology and Trades, Engineering, Math), but a much bigger percentage (32.35 percent) switch to Social Sciences and Business Administration (SSBA) majors. The students that switch to a major not within the bigger groupings, usually have to start the new program from year one because the course content is very different, and so most of the course work completed in their first year PSE cannot be transferred to the new program.

In Canada, the government subsidizes more than half of the tuition in post-secondary institutions. Considering the province of Ontario only, for the starting year in an arts major the provincial government invests CAD 3,100.00 per student in a university program and CAD 4,400.00 per student in a college program. The amount is higher for business and technology majors, and much higher for the upper university years.

The investment on the students that drop out after first year PSE, is a loss in net resources. The same is valid for the subsidy invested in those students that choose to switch to a different major and have to start a new program from the beginning. In Ontario only, on average this costs to the government at least 37 million Canadian dollars per year. The overall cost in the economy (government plus the private individual cost) amounts to at least 52 million canadian dollars.[1] Although individuals still may learn something useful (besides that they did not like the program) and improve some aspect of their cognitive skills by taking a single year in a program, yet the resources spent are not trivial and not used to the full potential. Obviously, it points to an inefficiency in the economy. It reflects an information friction, and the society would be better off if students had better informa-

[1]Author's calculations using the data posted in the Ministry of Training, Colleges and Universities and Statistics Canada. I add together the proportion of students in the YITS that drop out and switch program across the major groupings of field of study as classified in Table 3.1. Then, I multiply this proportion with the total number of 19 year old students in Canada that started a post-secondary program in 2004 in a university and college, separately. This results in the total number of students who disrupted their PSE path after their first year in PSE. This number is multiplied by the minimum amount of subsidy that the government of Ontario invests on each new student in their first year of university (CAD 3,100.00) and college (CAD 4,400.00).

tion about their own abilities before enrolling. This is the main reason why this topic is worthy of more attention than it has attracted so far.

In order to identify ways that lead to a minimum rate of inefficiency in PSE investment, one should look deeper into the reasons that lead students to these choices in the first place, which they end up fixing later at a cost to the government and themselves. Some relate the change in PSE path (dropping out or changing program) to the uncertainty on whether the student will graduate from the PSE program that they start (Hussey and Swinton, 2011; Altonji, 1993), which depends mainly on the quality of the match between students' own ability and major. The students form expectations on their probability to graduate from the program only after being introduced to the course content and experiencing the difficulty of the program through the first few semesters in PSE. Manski and Wise (1983) see PSE as part of a search process and state that students "*may derive informational value from attendance, even if they drop out*". Building on this idea, Manski (1989) constructs a model of PSE enrolment and completion which treats PSE schooling as experimentation and emphasizes the role of the ex ante probability of completion. Then Manski (2004) raises the importance of studying expectation formation using subjective expectations and beliefs elicited directly from the respondent. A few studies analyze expectation formation about social security benefits (Bernheim, 1988), expected earnings (Dominitz, 1998), returns to schooling (Dominitz and Manski, 1996) and the choice of contraception method (Delavande, 2008). Two recent papers apply a similar idea to the expectation formation during PSE. Zafar (2011) analyzes how students change their beliefs about post secondary outcomes and finds that students update their beliefs based on the unexpected content of information they receive from the change in their grade point averages, i.e. how well they do in the first year of their program. Similarly, Stinebrickner and Stinebrickner (2012) confirms this result through a higher frequency survey data analysis and also finds evidence that the student's probability to drop out changes significantly as the students learn about their academic ability. Earlier studies use choice data to infer about the way individuals form expectations. Both, Zafar (2011) and Stinebrickner and Stinebrickner (2012) rely on the elicited subjective beliefs in probabilistic form that enable the authors to avoid the use of an expectations formation model framed by strong assumptions. Both papers are able to identify a source of unexpected information, which is the driving force in belief updating.

This paper is similar in spirit to Zafar (2011) and Stinebrickner and Stinebrickner (2012). While they are among the first to employ subjective elicited beliefs in the educational choice setting, there are some limitations in the datasets that they use. They rely on survey data collected on one American post-secondary institution (university or college), small sample sizes and have little (or no) information about the students before entering the institution. This limits their ability to account for the heterogeneity among individuals. Both Zafar (2011) and Stinebrickner and Stinebrickner (2012) do not follow the survey respondents unless they stay within the institutions where the survey is conducted. This leads to an overestimate in the proportion of students that drop out and to an underestimate of the proportion of those who switch program. Additionally, there may be direct and indirect reasons why expectations lead to a disruption in the PSE path. For the direct reason, one has to repeat courses or maybe is required to leave university permanently or for one year if their grades are too low. This is a mechanical reason and not related to expectations but simply due to low grades. The indirect reason is due to students thinking after an unexpected drop in the GPA that they are in the wrong program and so switch programs even though they meet the progression requirement in the program they have started. Zafar (2011) and Stinebrickner and Stinebrickner (2012) can not separate the two reasons.

Different from these two studies, the analysis in this paper is based on a dataset which provides a wealth of information on the Canadian youth from all of the educational institutions in Canada and thus provides big sample advantages. The data contains information on the students' ability, family and high school. The longitudinal nature of the data allows tracking the students across programs and institutions. I am able to separate among those students that were required to leave the program because of a poor performance during the year. This group were only 2.1 percent of the sample, and are excluded from the analysis. The main specification is conditioned on a set of predetermined variables regarding individual and socio-economic background, a measure of cognitive ability and a measure for effort dedicated to studies during the first year in PSE. I estimate the effect that the surprising information has on updating several outcomes related to PSE and the labour market through years. The investigated outcomes are (i) the highest level of education one thinks s/he will attain, (ii) the probability to switch a major or program, (iii) to drop out of PSE and (iv) to switch institution (for instance from university to college).

The results indicate that a drop in the first year grade point average (GPA) of the students, relative to their overall high school GPA, is one main determinant affecting the revisions of expectations about the highest level of education students believe they will attain. The unexpected change in GPA is treated as a measure of new information that helps students update their perceptions of their own ability to access and complete learning activities in a specific major and institution. The probability of a disruption in the PSE path (dropout, switch program) increases if the students experience an unexpected decrease in GPA. The nature of the unexpected information seems to matter also; students' expectations and their likelihood of a change in their PSE path are only responsive to unexpected decreases in GPA, not positive ones. One explanation for this, may be that the decreases in GPA may be viewed more persistent than increases.

Appropriate policy measures may be taken to minimize the proportion of students that decide to drop out and change program in PSE. One way could be to provide consultation to high school students and inform them about the course content and difficulty level of the program they are about to enrol. One-on-one help regarding the overlap of qualifications they need for their dream job, their field of study preferences and the ability to well in certain majors may be more costly but more effective.

The paper is organized as follows. In section 3.2 the data and sampling characteristics are described. Section 3.3 summarizes the model that serves as a basis for the main specification and states the empirical results regarding the updating process in expectations. Section 3.4 investigates how decisions change with learning. Section 3.5 concludes.

3.2 Data

The Youth in Transition Survey was conducted at the same time on two cohorts of different age. The younger cohort is followed from 15 years old till 25. I am interested in their choices when 19 years old. At this age more than 60 percent of the students have started their education in a post-secondary institution. The analysis in this paper is based mainly on the younger cohort, but data from the older cohort are used in order to complement the analysis. This cohort was first surveyed when 18-20 years old, i.e. each student is 3 to 5 years older than the students in the younger cohort.

The sample size used in this empirical paper is composed of 7659 observations, 57 percent of which are female, 95.5 percent are single, and 95.7 percent are Canadian citizens. In this group, 26 percent first started their PSE studies in a college and 65 percent started a program in a university. After experiencing their first year in PSE, 38 percent of the college students switch to a university program and 20 percent of the university students switch to a college program. About 15 percent of the students switch major and two thirds of them switch major within the same institution.

The main interest of this paper is to find an explanation of PSE path disruptions. Theory suggests to look closer at how students form expectations and that their grades are the main determinant of persistence in PSE. Table 3.2 shows the student GPA distribution.

High school GPA is elicited from the following question: “In your last year of high school, (junior high or elementary school), what was your overall grade average, as a per-

Table 3.2: High school overall GPA and first year PSE GPA

	High School GPA_{HS}	First year PSE GPA_{PSE}
Younger Cohort		
90% or above	14.92	6.69
80-89%	44.40	26.52
70-79%	33.89	41.71
60-69%	6.01	20.23
50-59%	0.49	3.83
Under 50%	0.29	1.03
ΔGPA =Change in GPA for the Younger Cohort		
<0	55	
0	32	
>0	12	
$E(\Delta GPA)$ =Average Change in GPA in the Older Cohort by major, by institution		
<0	65	
0	12	
>0	23	
$\Delta GPA - E(\Delta GPA)$		
<0	55.42	
0	0.16	
>0	44.42	
Observations	7659	

Note: Given the categorical nature of the variable, each entry is a percentage of the respondents that report each level of high school and first year PSE GPA and the corresponding differences.

centage?". The GPA in the first year PSE program is retrieved from the following question: "In your first year, what was your overall grade average as a percentage?". The students were asked to check one of the 10 grade point categories listed in Table 3.2. Note that only 32 percent of the students manage to keep their grades within the 10 grade category of their high school GPA, 55 percent experience a decrease in the first year PSE and only 12 percent experience an increase in GPA.

PSE plans are elicited from the following question: "As things stand now, what is the highest level of education you think you will get?". The students were asked to choose one of the 10 levels of education: Some High School or less, High school diploma, Some post-secondary education level courses (no diploma or degree), Private business school or commercial school diploma; College, CEGEP, or trade/vocational certificate or diploma;

University degree or certificate below Bachelor's degree; University Bachelor's degree; University first professional degree (medicine, dentistry, veterinary medicine, law, optometry, divinity); Master's degree or University graduate diploma or certificate (above Bachelor's degree); PhD (or other earned doctorate, D.Sc., D.Ed.). Table 3.3 contains a tabulation of responses aggregated to fewer categories.

Table 3.3: Highest level of education that think will get in the future

	Expectations at 17	Expectations at 19
Less than College Diploma	5.22	3.55
College Diploma	18.41	21.33
Bachelor's Degree	49.40	47.37
Professional Degree	7.13	4.77
Master's Degree	14.57	17.94
PhD	5.27	5.04
Observations	7659	

Note: Given the categorical nature of the variable, each entry is a percentage of the respondents that report the expected level of education.

After their first year in PSE (age 19), a higher percentage expect to complete a college diploma, a Bachelor's degree and a Master's degree. Fewer students expect to attain a professional degree or a PhD degree. So, there can be observed some changes in the expectations after the students go through the first few semesters in their PSE program.

Since students with different backgrounds and abilities update their expectations differently, without accounting for their predetermined characteristics we may be facing an omitted variable bias in estimates. To prevent this, a set of control variables is incorporated in the main specifications. These include the PISA score which is a standardized measure for cognitive ability, a set of socio-economic variables and a set of other variables that describe the students' experience during their first year in PSE. The PISA score has a central role in accounting for the heterogeneity among respondents. In the sample, the score varies between 909.52 and 120.60 with a mean (standard deviation) of 557.85 (84.97). Among the major switchers the mean (standard deviation) for PISA score is actually above the sample average and has lower variation, 575.74 (78.70), and those who drop out have a lower PISA score on average, 542.55 (81.79), with higher variation than the overall sample.

The set of PSE outcomes that I analyze in this paper are the probability to drop out, switch major and switch program level (college vs. trades vs. university). These variables take a value of one to indicate the choice accordingly, and zero if the student makes no change in choices.

3.3 Expectations Updating about Highest Level of Education

In modelling the choice of field of study, recent papers have shown that accounting for the uncertainty to complete a program affects how students form expectations and, consequently, how they make choices (Altonji, 1993; Arcidiacono et al., 2011). In YITS, 17 year olds are on average 81 percent certain that they will eventually have the job they are interested in when 30 years old. Since the PSE program that they complete determines by and large their career path, the uncertainty that the students report is partly related to the possibility of being unable to complete the requirements and graduate from the PSE program that they start. When 21 years old, and at least a year before graduating from PSE, their certainty level increases to only 82.25 percent regarding this outcome.[2] Obviously, uncertainty follows the students throughout the studies until graduation, but decreases as they prove to themselves that they are able to complete the program. Thus, as they successfully progress in their program, their uncertainty decreases and they update their expectations on the future outcomes accordingly.

In this section I look into exactly that, i.e. I analyze how students update their expectations about a future outcome as they learn about their own academic ability. The outcome of interest is the highest level of education that they think they will get. Bernheim (1988) constructs an expectations updating model about Social Security Benefits during the pre-retirement period. The outcome of interest is the expectation about the value of an event X that will be realized at some point in the future. Suppose the individual i at time t has expectation X_{it}^e on the expected value of event X. At time t the individual is exposed to information Ω_{it}. Between time t and $t+1$, the individual is exposed to some additional new information, ω_{it+1}. Assuming that the individual remembers Ω_{it}, then the information content he is exposed to at time $t+1$ can be represented as $\Omega_{it+1} = (\Omega_{it}, \omega_{it+1})$. In notation,

$$X_{it}^e = E(X|\Omega_{it})$$

[2]Referring to the job or the type of business the students reported to be interested in having when they are 30 years old, the students were asked "How certain are you that you will eventually have this career or work?". The possible responses were: Very certain (coded as 100 percent certain), fairly certain (coded as 75 percent certain), fairly uncertain (coded as 50 percent certain) and very uncertain (coded as 25 percent certain).

Similarly,

$$X^e_{it+1} = E(X|\Omega_{it+1}) = E(X|\Omega_{it}, \omega_{it+1})$$

Then we can write,

$$E(X^e_{it+1}|\Omega_{it}) = E(E(X|\Omega_{it}, \omega_{it+1})|\Omega_{it}) = E(X|\Omega_{it}) = X^e_{it}$$

which implies that:

$$X^e_{it+1} = X^e_{it} + \eta_{it+1} \text{ where } \quad E(\eta_{it+1}|\Omega_{it}) = 0 \tag{3.1}$$

η_{it+1} is a function,$\psi(.)$, of the surprises received between time t and $t+1$. The unexpected new information, is defined as new information, ω_{it+1}, minus the expected portion of the new information, $E(\omega_{it+1}|\Omega_{it})$. Formally, we may write η_{it+1} as:

$$\eta_{it+1} = \psi\left[\omega_{it+1} - E(\omega_{it+1}|\Omega_{it})\right] \tag{3.2}$$

Substitute equation (3.2) in to (3.1) to get:

$$X^e_{it+1} - X^e_{it} = \psi\left[\omega_{it+1} - E(\omega_{it+1}|\Omega_{it})\right] \tag{3.3}$$

which identifies changes in expectations on the value of a future event as a function of unexpected new information that the individual encounters between time t and $t+1$.

If individuals are rational and use all available information in forming expectations, then any revision in their expectations will be determined by new information shocks. Zafar (2011) adopts this model to the expectations updating on academic and labour market outcomes. The same model is used in the curent paper, it serves as a basis for the empirical specification.

The outcome of interest, i.e. X, is the highest level of education that an individual will attain in his/her life. X_{it}^{e} and X_{it+1}^{e} are the students' responses two years apart to the following question: *"As things stand now, what is the highest level of education you think you will get?"*. The question is carefully formulated to retrieve $E(X|\Omega_{it})$ from each student during the Winter 2002 and 2004 semesters. I identify the source of new information as the unexpected portion of the change (increase/decrease) in the GPA from high school to PSE. The students were asked about their high school GPA realized before December 2001. Then, they were interviewed in Winter 2004 and asked to report the first year PSE grade average which was realized in Winter 2003. Unfortunately, the survey did not ask the students about how much they expect their GPA to increase or decrease in the first year in PSE when compared to their high school GPA. This is because the difference between the realized change in GPA and the expected change in GPA yields the unexpected portion of the change in GPA, and also the source of new unexpected information. Unlike previous studies, that use the change in GPA within the first few semesters/years in PSE, by using the change in GPA between high school and PSE I am capturing the initial learning about own ability that students experience in PSE studies.

Referring to equation (3.3) above, ω_{it+1} in this case is the grade change that the younger cohort experiences $\Delta GPA = GPA_{PSE} - GPA_{HS}$; $E(\omega_{it+1}|\Omega_{it})$ is the grade change that the older cohort experiences serving as a proxy for $E(\Delta GPA|\Omega_{it})$. Note that the distinction between and ω_{it+1} and $(\omega_{it+1} - E(\omega_{it+1}|\Omega_{it}))$ is crucial. To explain this point, suppose we were to interpret the change in realized GPA, i.e. ΔGPA, as new information. Then we may confound a positive piece of new information with a negative one if the first year PSE grades decreased by less than the student expected. This highlights the importance of expectations about the change in GPA that is not available in YITS. However, I have another way to deal with this data limitation. If economic agents form rational expectations and use all available information, the most important piece of information that students would use to form expectations about their change in grades may be the change in grades that an older cohort experienced in the same PSE institution (university or college) and major. So, I assume that the older cohort's change in GPA between high school and first year PSE is a good proxy for the younger cohort's $E(\omega_{it+1}|\Omega_{it})$. Because I can observe the PSE institution for both cohorts and their major, I am able to calculate an average change in GPA by institution and by major and merge it with the data of the younger cohort.

I start by assuming a linear functional form for $\psi(.)$. The main specification is equation 3.4.

$$X^e_{it+1} - X^e_{it} = \psi\left[\omega_{it+1} - E(\omega_{it+1}|\Omega_{it})\right] + Z'_{it}\gamma + \epsilon_{it+1} \tag{3.4}$$

where Z_i includes: PISA score, gender indicator, marital status indicator, citizenship indicator, province of residence dummies, indicator if student receives monetary transfers from parents, parental expectations, parental education. I also condition on a set of other variables that describe the students' experience during their first year in the PSE program. In this way I can isolate the effect of the change in grades from motivation, effort, willingness to adjust, information value earned from that one year in PSE. These variables include: "Sure about future type of work", "Workshops to adjust", "Hours of homework", "Times though of dropping out", "Right program", "Idea for future plans", "Job market skills". Detailed definitions of the variables are included in the Appendix.

Table 3.4 contains the OLS estimation results of equation (3.4). I consider two separate regressions, column (1) and (2), which differ in the independent variable of interest. In column (1), I use new information as the variable that captures learning about own ability. In column (2) instead, I use unexpected new information as the learning variable. In both cases the coefficients are positive and statistically significant suggesting an increase of 0.08 in the level of education if the student's GPA increases by an unexpected 10 percent. Note that the coefficient is higher in the first column when the expected change in GPA is not subtracted, suggesting a bias.

Table 3.4: Dependent Variable: Δ Expected Highest Education Level to Attain

	(1) $\Delta GPA = GPA_{PSE} - GPA_{HS}$	(2) $\Delta GPA - E(\Delta GPA)$
Coefficient	0.084***	0.080***
	(0.021)	(0.021)
Pseudo R^2	0.016	0.016
Sample Size	5594	5594

Note: ***Significance at 1%, **Significance at 5%, *Significance at 10%. Standard errors in parenthesis.

The dependent variable in Table 3.4 is categorical and takes count values between -7 and 7. The histogram of this variable resembles more a normal distribution than to a Poisson and Negative Binomial distribution. The Bayesian Information Criterion was smallest for the OLS regression when compared to the Poisson and Negative Binomial models. For these reasons I chose not to apply a count model estimator.

Next I relax the linearity assumption in the functional form for $\psi(.)$. I use non-parametric methods to estimate the relationship between the update in expectations and the information shocks. See Figure 3.1 for the local polynomial regressions estimated with degree three and Epanechnikov kernel and Silverman's bandwidth. The figure contains the graphs, the first and the second are regression estimates using new information and unexpected new information as the variable of interest, respectively. The third graph contains both graphs together for comparison purposes. Different from the story that OLS estimates suggest, in figure 3.1 notice that the PSE expectations updating is responsive only to unexpected decreases in GPA. Students do not react to positive changes in GPA in updating their PSE expectations. Also, note that, as expected, the response is much bigger when I use the unexpected new information. This is reflected in the steeper slope of the local polynomial regression (solid line) in the third panel of the figure.

Dominitz (1998) finds a similar result using only the “new information” variable. One explanation he gives for this result, which may apply in this paper’s setting too, is that a negative change in the independent variable of interest may be perceived as more persistent than a positive change.

3.4 PSE Choices as Response to New Information

In the YITS, when asked *“What was the main reason you chose to take (this program/-subject)?”*, 68 percent of the students in YITS reported that they had self interest on the major and 16 percent reported reasons related to future job/earnings. So the major factor in choosing a field of study is in fact “personal interest”. Students form expectations about the course content of the PSE program they choose to pursue based partly on their high school curricula, which may not be sufficient especially for some majors. During the first year in PSE, some of them realize that the program is not suitable and leave PSE or change program. In the YITS, more than 18 percent of the students who started a post-secondary education (PSE) program in the 1998-1999 academic year, report to have changed program after the first year in PSE, and about 10 percent report to have switched major within institution. When asked *“What is the main reason you changed it?”* (referring to the main field of study or specialization), 41 percent of the students responded by *“Didn’t like it/not for*

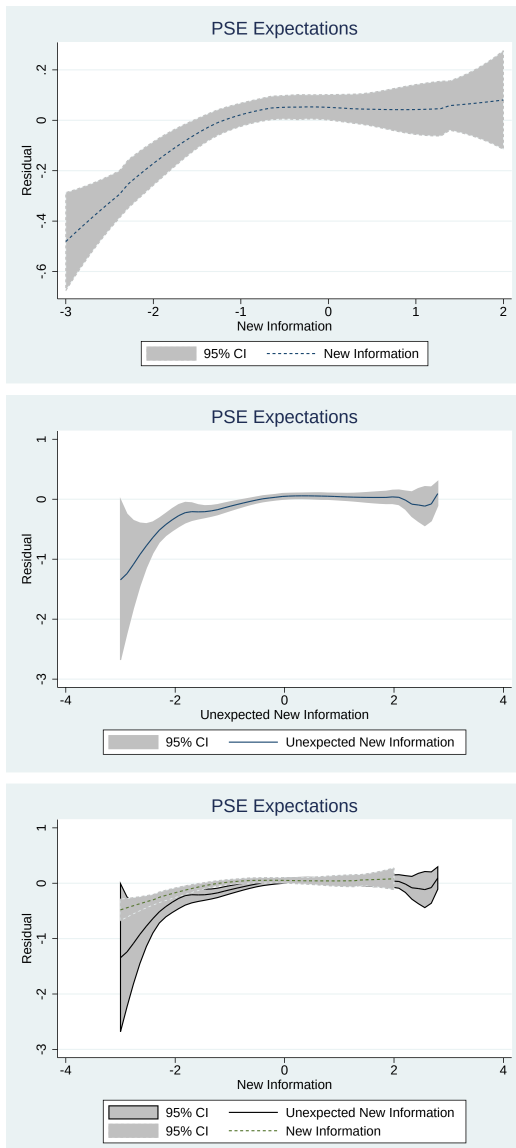


Figure 3.1: Expectations About Highest Level of Education and New Information - Local Polynomial Regression

me" and 38 percent of the students answered by *"Interest in new subject"*.

Stinebrickner and Stinebrickner (2012) conduct a longitudinal survey on low-income family students. The data retrieved from a carefully constructed survey enables them to see further into the trajectory of the students' choices. They find evidence that the students choose to change their PSE path as they learn about their academic ability in that program through their grade performance. I use a similar specification to see how students change decisions. The main specification is equation (3.5).

$$P(Y_{it+1} = 1) = \psi \left[\omega_{it+1} - E(\omega_{it+1}|\Omega_{it})\right] + \gamma Z_i + \epsilon_{it+1} \tag{3.5}$$

where Y is the outcome, Z_i the predetermined variables for each respondent i and ϵ_{it+1} is a normally distributed $N(0, \sigma^2)$ error term. The outcomes of interest in this section are the probability to switch a major, probability to drop out and the probability to change institution level (e.g. from college to university). The results are summarized in Table 3.5 for the case when the linear functional form assumption is kept. Using either the "new information" variable or the "unexpected new information" variable the results are very similar and suggest a decrease in the probability of dropping out and switching major or institution by 1 percent if there is an unexpected increase in the realization of GPA. For an unexpected increase of 10 grade points, the probability to switch major decreases by 1.8 percent, the probability to switch from college to university or vice versa decreases by 0.3 percentage points and the probability to drop out of PSE decreases by about 1 percentage point. The unexpected change in GPA seems to have the highest impact on the probability to switch major.

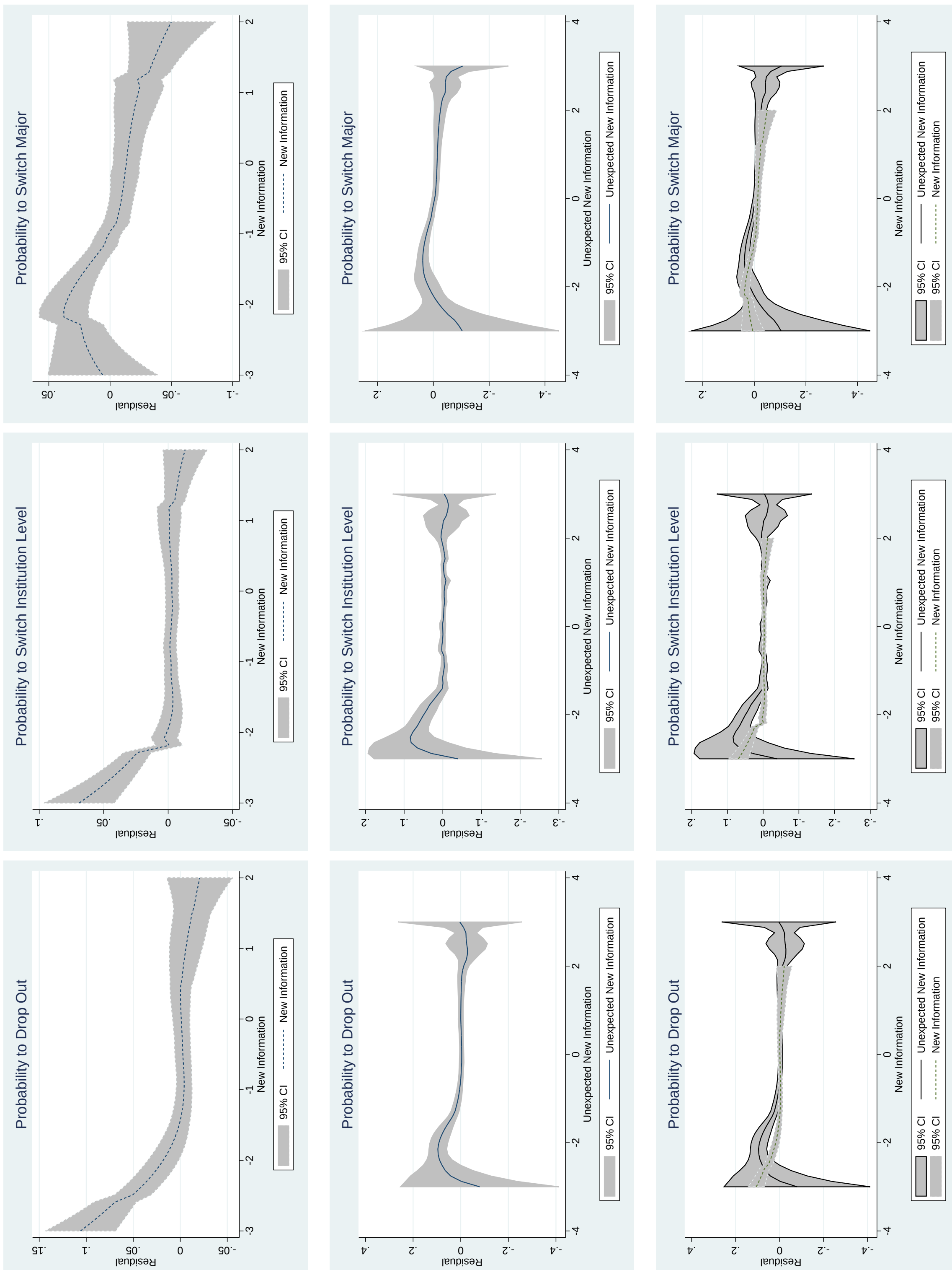
Probit regressions, however, provide an average marginal change in the probability of the outcome of an increase by 10 grade points. Next I show that the marginal effects are in fact very different for different values of the change in grades. So, I allow $\psi(.)$ to follow a non-linear function estimate equation 3.4 as a semi-parametric Partial Linear Model regression. The estimation is performed in two steps. First I extract the projection of $P(Y = 1)$ on Z and retrieve the unexplained part or residuals. Then, I fit a local polynomial equation of degree three to the projection of residuals on the unexpected change in GPA. Figure 3.2 contains the results. Each column of graphs contains the local polynomial regression estimated separately by using the two different independent variables. For reasons stated earlier in the paper I concentrate on the middle row of graphs. Obviously, the results are very different from those of Table 3.5. For each of the outcomes, the probability to drop out of PSE, to change major or institution is responsive to unexpected decreases in the students' GPA only and within a range of unexpected decreases between 10 to 20 grade points. As expected, the probability of each outcome decreases as the unexpected decrease in GPA becomes smaller in magnitude. When comparing the graphs that use simply the difference between the GPA realizations to the ones that use unexpected changes in GPA as independent variable (last row in figure 3.2), notice that the former leads to a misleading result.

Table 3.5: New information and probability to change choices

Dependent Var	$\hat{\beta}$	Pseudo R^2	Sample Size
Panel A - New Information			
Pr(Switch Major)	-0.018*** (0.004)	0.125	5838
Pr(Switch Institution Level)	-0.003*** (0.001)	0.206	5150
Pr(PSE Drop-out)	-0.010*** (0.002)	0.207	5847
Panel B - Unexpected New Information			
Pr(Switch Major)	-0.017*** (0.004)	0.125	5838
Pr(Switch Institution Level)	-0.003*** (0.001)	0.205	5150
Pr(PSE Drop-out)	-0.009*** (0.002)	0.206	5847

Note: ***Significance at 1%, **Significance at 5%, *Significance at 10%. The control variables include Female dummy, age, single dummy, citizen dummy, parental education, financial help from parents dummy, aspirations for PSE, motivation to achieve desired level of education, rate of skills on math, writing, computer, reading, oral communication, solving problems, volunteering activities indicator, rural residence indicator, province indicator.

Figure 3.2: Probability of a Choice and New Information - Local Polynomial Regression



3.5 Conclusion

In this study I follow the new literature on expectations formation about post-secondary education outcomes which explain the updating process by the unexpected new information that the individuals receive in interim. Earlier work has used data on surveys conducted in a single institution and thus, very small sample sizes. The data I am using allows us to conduct a large scale study by using data from students from the majority of Canadian post-secondary institutions. I find evidence that expectation revisions are triggered by unexpected information. In this setting the unexpected information is measured by the unexpected change in grade point average from high school to post secondary education. The bigger the unexpected decrease, the higher is the probability to drop out of PSE and to switch to another program.

Differently from previous work on this topic, I consider the possibility that the effect of unexpected shock may not be even and independent of the magnitude of the shock. I use non-parametric methods to estimate a non-linear relationship. The expectation revisions and choices are only responsive to "unexpected bad news". By "unexpected bad news" I mean that as students receive a unexpected decrease in GPA, this is perceived as a unfit of their ability to do well in the PSE program, so they decide to switch to another major or drop out.

Disruptions in the PSE path lead to ineffective government investment in PSE. Appropriate policy measures may be fitted to minimize this inefficiency in the system. One way could be providing consultation to high school students and informing them about the course content and difficulty level of the program they are about to enrol. This could potentially improve the student ability-to-major match.

Appendix: Variable Definitions

Dependent Variables

Expected highest level of education: Categorical variable varying from 1 to 10 as the responses to the question *"As things stand now, what is the highest level of education you think you will get?"* vary between the following increasing categories: *Some High School or less*, *High school diploma*, *Some post-secondary education level courses (no diploma or degree)*, *Private business school or commercial school diploma*, *College, CEGEP, or trade/vocational certificate or diploma*, *University degree or certificate below Bachelor's degree*, *University Bachelor's degree*, *University first professional degree (medicine, dentistry, veterinary medicine, law, optometry, divinity)*, *Master's degree or University graduate diploma or certificate (above Bachelor's degree)*, *PhD (or other earned doctorate, D.Sc., D.Ed.)*.

PSE drop out: Dummy variable equals to 1 if student reported to have dropped out of PSE after their first year in PSE.

Switch major: Dummy variable equals to 1 if student reported to have switched major after their first year in PSE.

Switch institution level: Dummy variable equals to 1 if student changed institution, e.g. from college to university or vice-versa, after first year PSE.

Independent Variables of Interest

ΔGPA $=$ GPA$_{\text{PSE}}$ $-$ GPA$_{\text{HS}}$: The difference between the GPA in the first year at a post-secondary institution and the overall high school GPA. Each variable takes categorical values of 1 to 6 for the options listed, respectively: *Under 50%*, *50-59%*, *60-69%*, *70-79%*, *80-89%*, *90% or above*.

ΔGPA $-$ E(ΔGPA): The difference between the experienced change in GPA between the first year at a post-secondary institution and the overall high school GPA, ΔGPA, and the expected change in GPA, $E(\Delta GPA)$. To measure $E(\Delta GPA)$ I use as proxy the change in GPA that the older cohort experienced in the same institution (university or college) and same field of study.

Control Variables

Female indicator: Dummy variable equaling 1 if respondent's gender is female, zero otherwise.

Single indicator: Dummy variable equaling 1 if respondent's marital status is single, zero otherwise.

Canadian citizen: Dummy variable equaling 1 if respondent is a Canadian citizen, zero otherwise.

Province indicators: Indicator variables for each of the provinces. Excluded category is Ontario.

Rural vs. Urban indicator: Dummy variable equaling 1 if respondent resides in a rural area, far commuting distance from the city centre.

Parental monetary transfers: Amount of monetary transfers that the student receives from their parents during first year PSE.

Parental expectations: Dummy variable equaling 1 if the respondent's parents respond that they expect at least one university degree from their child, zero otherwise.

Parental education: Dummy variable equaling 1 if parent has a post-secondary education degree. This variable is separate for mother and father.

PISA score: Control Variables Programme for International Student Assessment (PISA) reading test score.

Sure for type of work: Categorical variable equaling 1 if response to question *"During my first year, I was sure of the type of work I would like to have in the future"* is *Strongly Disagree*, 2 if response is *Disagree*, 3 if response is *Agree* and 4 if response is *Strongly Agree*.

Workshops to adjust: Dummy variable equaling 1 if response to question *"During or before your first year, did you take part in any workshops, programs or courses designed to help you adjust to first-year studies?"* is 1 if response is *Yes* and zero otherwise.

Hours homework: Categorical variable varying from zero to 7 as the responses to the question *"During your first year, about how many hours each week did you spend studying or doing assigned work outside of class?"* vary between the following increasing categories: *Zero*, *Less than one hour per week*, *1 to 3 hours*, *4 to 7 hours*, *8 to 14 hours*, *15 to 20 hours*, *21 to 30 hours* and *More than 30 hours per week*.

Times thought of dropping out: Categorical variable varying from zero to 4 as the responses to the question *"How many times per month did you think about dropping out. Was it ...?"* vary between the following increasing categories: *Never*, *Less than once a month*, *Once or twice a month*, *About once a week*, *More than once a week*.

Right program: Categorical variable equaling 1 if response to question *"During my first year, I felt I had found the right program for me"* is *Strongly Disagree*, 2 if response is *Disagree*, 3 if response is *Agree* and 4 if response is *Strongly Agree*.

Idea for future plans: Categorical variable equaling 1 if response to question *"During my first year, helped me get a better idea of my future plans"* is *Strongly Disagree*, 2 if response is *Disagree*, 3 if response is *Agree* and 4 if response is *Strongly Agree*.

Job market skills: Categorical variable equaling 1 if response to question *"My first year, gave me skills that help me in the job market"* is *Strongly Disagree*, 2 if response is *Disagree*, 3 if response is *Agree* and 4 if response is *Strongly Agree*.

References

Alberto Abadie and Guido W. Imbens. Simple and bias-corrected matching estimators. *Technical report, Department of Economics, University of California, Berkeley.*, 2002.

Alberto Abadie and Guido W. Imbens. Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1):235–267, 2006.

Joseph G. Altonji. The demand for and return to education when education outcomes are uncertain. *Journal of Labor Economics, Part 1: Essays in Honor of Jacob Mincer*, 11 (1):48–83, 1993.

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.

Joshua D. Angrist and Jorn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2008.

Peter Arcidiacono. Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1-2):343–375, 2004.

Peter Arcidiacono, V. Joseph Hotz, and Songman Kang. Modeling College Major Choices Using Elicited Measures of Expectations and Counterfactuals. *SSRN eLibrary*, 2011.

B. Douglas Bernheim. How do the elderly form expectations? an analysis of responses to new information. *NBER Working Paper Series, Working Paper No. 2719, National Bureau of Economic Research, Inc.*, 1988.

Julian Betts, Christofer Ferrall, and Ross Finnie. The role of university characteristics in determining post-graduation outcomes: Panel evidence from three recent canadian cohorts. *Analytical Studies, Research Paper Series, Statistics Canada, Catalogue No. 292*, 2007.

Michela Bia and Alessandra Mattei. A stata package for the estimation of the dose-response function through adjustment for the generalized propensity score. *The Stata Journal*, 8 (3):354–373, 2008.

Dan Black, Kermit Daniel, and Jeffrey Smith. College quality and the wages in the United States. *German Economic Review*, 6(3):415–443, 2005.

Dan A. Black and Jeffrey A. Smith. How robust is the evidence on the effects of college quality? evidence from matching. *Journal of Econometrics*, 121:99–124, 2004.

Dan A. Black and Jeffrey A. Smith. Estimating the returns to college quality with multiple proxies for quality. *Journal of Labour Economics*, 24 (3):701–728, 2006.

Sandra E. Black, Paul J. Devereux, and Kjell G. Salvanes. Under pressure? The effects of peers on outcomes of young adults. *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. *Institute for the Study of Labor, IZA Discussion Paper No. 4723*, 2010.

Dominic J. Brewer, Eric R. Ride, and Ronald G. Ehrenberg. Does it pay to attend an elite private college? *Journal of Human Resources*, pages 106–123, 1998.

David Card. The causal effects of education on earnings. In O. Ashenfelter and David Card, editors, *Handbook of Labour Economics*, volume 3, chapter 30, pages 1801–1863. Elsevier Science B.V, 1999.

Arnaud Chevalier and Gavan Conlon. Does it pay to attend a prestigious university? *Center for Economics of Education, London School of Economics and Political Science, Houghton Street, London WC2A 2AE*, 2003.

Louis N. Christofides, Michael Hoy, Zhi (Jane) Li, and Thanasis Stengos. The evolution of aspirations for university attendance. In Ross Finnie, 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.

W. Cochran and Donald B. Rubin. Controlling bias in observational studies. *Sankyha*, 35: 417–446, 1973.

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

Stacy Dale and Alan B. Krueger. Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics*, 79(2):1491–1527, 2002.

Kathleen Day. The effect of high school resources on investment in post-secondary education in Canada. *Unpublished Manuscript*, 2009.

R. H. Dehejia and S. Wahba. Causal effects in non-experimental studies: Re-evaluating the evaluation of training programmes. *Journal of the American Statistical Association*, 94: 1053–1062, 1999.

Adeline Delavande. Pill, patch or shot? subjective expectations and birth control choice. *International Economic Review*, 49(3):999–1042, 2008.

Weili Ding and Steven F. Lehrer. Do peers affect student achievement in China's secondary schools? *The Review of Economics and Statistics*, 89(2):300–312, May 2007.

Jeff Dominitz. Earning expectations, revisions, and realizations. *Review of Economics and Statistics*, 80:374–88, 1998.

Jeff Dominitz and Charles F. Manski. Eliciting student expectations of the returns to schooling. *Journal of Human Resources*, 31(1):1–26, 1996.

Kent Eliasson. The role of ability in estimating the returns to college choice: New Swedish evidence. *Umea Economic Studies No. 691. Umea University.*, 2006.

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 Finnie, 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 drop-out decision. *Institute for Fiscal Studies, IFS Working Paper No. W09/21*, September 2009.

Mark Frenette. Career goals in high school: Do students know what it takes to reach them and does it matter? *Statistics Canada, Social Analysis Division, Research Paper Series No. 320*, 2009.

Mark Frenette and Klarka Zeman. Why are most university students women? evidence based on academic performance, study habits and parental influences. *Statistics Canada, Business and Labour Market Analysis, Analytical Studies-Research Paper Series 11F0019 No. 303*, 2007.

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

Eric A. Hanushek and Ludger Woessman. The role of school improvement in economic development. *CESifo, Working Paper No. 1911, Category 5: Fiscal Policy, Macroeconomics and Growth*, 2007.

Eric A. Hanushek, John F. Kain, Jacob M. Markman, and Steven G. Rivkin. Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5):527–544, 2003.

James J. Heckman, H. Ichimura, and Petra E. Todd. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, 64:605–654, 1997.

James J. Heckman, H. Ichimura, J.A. Smith, and Petra E. Todd. Characterising selection bias using experimental data. *Econometrica*, 66(5), 1998a.

James J. Heckman, H. Ichimura, and Petra E. Todd. Matching as an econometric evaluation estimator. *Review of Economic Studies*, 65:261–294, 1998b.

James J. Heckman, R. J. LaLonde, and J. A. Smith. The economics and econometrics of active labour market programmes. In O. Ashenfelter and David Card, editors, *The Handbook of Labour Economics*, volume 3A. 1998c.

James J. Heckman, Lance J. Lochner, and Petra E. Todd. Fifty years of Mincer earnings regressions. *National Bureau of Economic Research, Inc., NBER Working Papers 9732*, 2003. URL `ftp://repec.iza.org/RePEc/Discussionpaper/dp775.pdf`.

Keisuke Hirano and Guido Imbens. The propensity score with continuous treatments. In *Missing Data and Bayesian Methods in Practice: Contributions by Donald Rubin's Statistical Family*. Wiley, 2004.

Linda Holmlund. The effect of college quality on earnings: Evidence from Sweden. *Department of Economics, Umea University. Mimeo*, 2009.

Julia Horstschraer. Quality signals in action? The importance of rankings and excellence labels for university choice of high-ability students. ZEW, Manuscript, August 2011.

Caroline Hoxby. Peer effects in the classroom: Learning from the gender and race variation. *NBER Working Papers No. 7867*, 2000.

Andrew J. Hussey and Omari H. Swinton. Estimating the ex ante expected returns to college. *American Economic Review: Papers & Proceedings*, 101(3):598–602, 2011.

Guido W. Imbens. The role of propensity score in estimating dose-response functions. *Biometrika*, 87(3):706–710, 2000.

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.

Estelle James, Nabeel Alsalam, Josef C. Conaty, and Duc-Le To. College quality and future earnings: Where should you send your child to college? *AEA Papers and Proceedings, Economics of Education Industry*, 79(2):247–252, 1989.

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.

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 achievement in english primary education. *IZA Discussion Paper No. 3866*, 2008.

Kevin Lang and Erez Siniver. Why is an elite undergraduate education valuable? Evidence from Israel. *Labour Economics*, 18(6):767–777, 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.

Edwin Leuven and Barbara Sianesi. Psmatch2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. *"http://ideas.repec.org/c/boc/bocode/s432001.html".This version 4.0.4 10nov2010 E. Leuven, B. Sianesi.*, 2003.

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`.

Mark C. Long. College quality and early adult outcomes. *Economics of Education Review*, 27:588–602, 2008.

Mark C. Long. Changes in the returns to education and college quality. *Economics of Education Review*, 29:338–347, 2010.

Charles F. Manski. Schooling as experimentation: a reappraisal of the postsecondary dropout phenomenon. *Economics of Education Review*, 8(4):305–312, 1989.

Charles F. Manski. Measuring expectations. *Econometrica*, 72(5):1329–1376, 2004.

Charles F. Manski and David A. Wise. *College choice in America*. Cambridge: Harvard University Press, 1983.

J. Mincer. Investment in human capital and personal income distribution. *Journal of Political Economy*, 66(4):281–302, 1958.

James Monks. The returns to individual and college characteristics, evidence from the national longitudinal survey of youth. *Economics of Education Review*, 19:279–289, 2000.

Philip Oreopoulos and Ryan Dunn. Information and college access: Evidence from a randomized field experiment. *NBER Working Paper 18551*, 2012.

Philip Oreopoulos and Kjell G. Salvanes. How large are returns to schooling? hint: Money isn't everything. *NBER Working Papers 15339, National Bureau of Economic Research, Inc.*, 2009.

Paul R. Rosenbaum and Donald B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

Paul R. Rosenbaum and Donald B. Rubin. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1):33–38, 1985.

Donald B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66:688–701, 1974.

Donald B. Rubin. Bias reduction using Mahalanobis-metric matching. *Biometrics*, 36: 293–298, 1980.

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

Todd R. Stinebrickner and Ralph Stinebrickner. Learning about academic ability and the college dropout decision. *Journal of Labor Economics*, 30(4):707–748, 2012.

Tuomo Suhonen. *Does university quality affect student's subsequent earnings? Evidence from Finland using field-of-study specific quality measures*. PhD thesis, School of Business and Economics, University of Jyvvaskyla, Finland, January 2011.

Jacob Vigdor and Thomas Nechyba. Peer effects in North Carolina public schools. *Unpulished*, 2005.

Basit Zafar. How do college students form expectations? *Journal of Labor Economics*, 29 (2):301–348, 2011.

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