I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11025

# The Context-Bound University Selectivity Premium

Joniada Milla

SEPTEMBER 2017

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11025

# The Context-Bound University Selectivity Premium

**Joniada Milla**
*Saint Mary's University and IZA*

SEPTEMBER 2017



**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

## The Context-Bound University Selectivity Premium*

In this paper I present a selective survey of the empirical literature on wage premium to university selectivity focusing mainly on the context of the country under analysis and the identification strategies employed. I then estimate the wage premium to university selectivity using Canadian data and two popular methods to correct for non-random selection in universities of different quality: matching methods and instrumental variables (IV). I estimate a wage premium of 7% using the matching estimator, and a premium of 14.8% using the IV estimator for alumni of selective Canadian universities 4–6 years after graduation. My findings are in line with the literature on countries with a moderately differentiated higher education system that has low variation in tuition fees and is well supported by public funds.

**JEL Classification:** C21, I23, J30

**Keywords:** university selectivity, wage premium, context

**Corresponding author:**
Joniada Milla
Department of Economics
Saint Mary's University
923 Robie St.
Halifax, NS B3H 3C2
Canada
E-mail: joniada.milla@smu.ca

* While the research and analysis are based on data from Statistics Canada, the opinions expressed do not represent the views of Statistics Canada. I am thankful to Thanasis Stengos, Miana Plesca, Michael Hoy, Adeline Delavande, Louis Christofides, Justin Smith, Mikal Skuterud and Elena Mattana for useful feedback on a previous draft.

# 1 Introduction

Labor market prospects for recent graduates depend mainly on their acquired skills, ability and work productivity. In particular, the last two are unobservable characteristics. From an employer's point of view, an inexpensive instrument that is often used as a signal for these unobserved attributes is the prestige, reputation or quality of the institution where the potential employee received their degree. Following the discussion in Broecke (2012), I will use the term university[1] selectivity in the rest of the paper. Hence, employers statistically discriminate based on two possible scenarios that point them to the better applicant for the job: either because better universities attract and skim the better students, or because better universities make better graduates (everything else equal). I survey the literature that recognizes and attempts to separate the two sources using different approaches, but this remains a challenging task given the lack of consensus in premium estimates. A non-contestable outcome of this survey is the differences in estimates with respect to the context of the higher education (HE) systems under evaluation.

The literature on university selectivity premium on wages is vast and focuses mainly on countries with a polarized education system (such as the U.S., Columbia or Chile), where the differences in university selectivity are evidently large. This paper revisits the question of wage premium to university quality in a different context, such as the Canadian one. The Canadian case is interesting and different from the countries that have been investigated in the literature. In Canada the most prestigious universities are public universities, relying at an almost equal extent on public and private funds. Although not yet shown analytically, sufficient evidence indicates that the resource allocation in Canada is relatively homogeneous, and in terms of student selectivity falls between the selective HE systems in the Americas (U.S., Chile, Colombia that have high tuition fees with a relatively high variance and select students based on a standardized entry exam), and the publicly funded HE systems in the E.U. that are equally endowed in terms of public resources, charge very low/no tuition with low variance and admit students based on their high school grade point average (GPA). In

[1] Differently from the United States, in Canada college and university refer to different academic entities. In Canada, colleges offer vocational trades programs and Bachelor-equivalent degrees in arts and sciences focused on practical aspect of learning. College instructors in Canada focus on teaching and have extensive experience in the private sector. Universities offer Bachelor degrees with research opportunities leading to graduate studies.

the later context, as expected the university selectivity premium is lower than in the former.

The main novelty of this paper is to provide the first study that addresses selection un unobservable characteristics in the context of Canada. The reason behind the lack of other research on the area is due to the absence of crucial information in Canadian surveys such as the name of the university that individuals completed. To the best of my knowledge, the only other paper that asks how university characteristics affect graduates' earnings for Canada is Betts et al. (2013). They provides a descriptive analysis using the Canadian National Graduates Survey, which while it contains information on the name of the university, lacks a measure for the individuals' academic ability. I use a new, previously unexploited data, the Canadian Youth in Transition Survey (YITS), which is the first Canadian data to contain both the above information and more: the high school grades as measure of ability, the name of the university that they enrolled and graduated from, and the province of residence at age 15, while a student at university and when last interviewed. This data also offers rich and detailed background information on individuals, their families, high school experience, and a measure of students performance during their first academic year in university (their grade point average). The YITS also contains information on generally unobserved characteristics such as aspirations, student-elicited rating of their skills on oral communication, writing and math. In addition, there is available information on the frequency at which the student volunteered, and the number of on- and off-campus clubs that the student is actively engaged in. I leverage the large set of variables in the YITS to correct for the non-random selection into universities of different degrees of selectivity based on these differences by using a matching estimator.

Although the YITS has an unusual set of information on university graduates, the possibility that other unobserved characteristics of individuals that guide selection into universities of different quality can not be ruled out unambiguously. To address this concern I use an instrumental variable estimator in line with Long (2008). The exclusion restriction this study uses, (namely, the average selectivity measure for the institutions within a geographic radius of the students high school residence) leads to a LATE (local average treatment effect) estimate reflecting mainly the wage premium of the subsample of students who live in regions with better universities, that in turns makes them more likely to attend a better university.

Consequently, the students that reallocate to attend a better university from regions that originally do not favor them in this respect, are in fact weighted relatively less. It is likely that the expected returns to university quality are higher for these students that are willing to invest more. I propose a second exclusion restriction that offers a new source of exogeneity and potentially leads to a more "inclusive" LATE estimate that is representative of both students that attend a university close to their parents' home or move out to a different province to do so. In addition to the Long (2008) instrument, I also use the difference of the average university selectivity between the students' pre-university residence region and the residence region while at university.

In this paper, I use all the three types of selectivity measures used in the literature: the average high school grade of the entering cohort (which is the main admission criteria in Canada), the "Best Overall Reputation Ranking" (hereafter Reputation Ranking) published yearly by the *Maclean's* magazine, and I also construct a quality measure using dimension reducing techniques in line with the previous literature. Using the distribution of the selectivity measures and the type of the university (research vs. teaching intensive) I group the universities into three groups from most to least selective: A, B and C. I find that graduates of selective Canadian universities (group A & B) earn 7% higher hourly wages than graduates of less selective universities (group C). These returns are not driven by observable differences. The instrumental variable estimates (around 8–14.8%), that are free of bias from unobserved characteristics, are higher in magnitude than the matching estimates, but the difference is statistically insignificant. Overall, the findings are in line with the literature on university selectivity that estimates a wage premium of 15–28% for graduates of the most selective universities in countries with a strongly differentiated HE system, and a premium of 2–7.4% in countries with moderately differentiated HE system.

The rest of the paper is organized as follows. I present a literature review in Section 2 while describing the Canadian HE system in comparison to that of other countries analyzed in the literature. I introduce the main specification and discuss the methodology in Section 3, the data in Section 4, and the results in Section 5. Section 6 concludes.

## 2 Literature Review

In this section I present a review of the literature on wage premium of university quality. All the papers discussed conduct the analysis of selectivity premium at the undergraduate level, their outcome of interest is the graduates' private wage rate, conduct a quantitative empirical analysis, acknowledge and attempt to address the non-random selection of students in universities of different quality.[2] In Table 1 I provide a tabulation of the wage premium estimates to university selectivity by estimation method and country.

**Estimation Method** In the early literature, the premium to university quality is most often explored using cross-sectional data in a "selection on observables" setting (Heckman and Robb, 1985).[3] The majority of these studies find a positive and significant premium of university quality on earnings. The recent literature has been more concerned with causal inference, and employs six different approaches (in addition to LS that is seen as the benchmark) to address the non-random selection of students in universities of different quality as shown in Table 1: (i) Matching Estimator (ME); (ii) sibling or twin Fixed Effects estimator (FE); (iii) Self-revelation (SR) model suggested by Dale and Krueger (2002); (iv) Applicant–matched (AM) model also suggested by Dale and Krueger (2002); (v) Instrumental Variables (IV) estimator as suggested in Long (2008); (vi) and the Regression Discontinuity Design (RDD) method introduced by Thistlewaite and Campbell (1960) and reviewed by Lee and Lemieux (2010) for the empirical economist.

As seen from Table 1, for the same country (and sometimes using the exact same data such as in Long (2008), Broecke (2012), Borgen (2014)) the different methods deliver different results. There are advantages to each of these methods, as there are weaknesses in their application. The main advantage of the matching estimator as argued in Black and Smith

[2]Even though excluded from this literature survey, it is important to note that the selectivity of graduate studies has been studied for the MBA program (Chen et al., 2012; Grove and Hussey, 2011), and the university selectivity at the undergraduate level has also been shown to affect other non-monetary outcomes such as dropping out behavior (Hanushek et al., 2008), college completion (Heil et al., 2014; Cohodes and Goodman, 2014), attending graduate school and expectations on wages (Loyalka et al., 2012), choice of siblings' PSE institution (Goodman et al., 2015), marriage market (Kaufmann et al., 2013), intergenerational effects (Kaufmann et al., 2014), PSE exit exam scores (Saavedra and Saavedra, 2011), and social returns such as becoming a more involved citizen through voting behavior (Solis, 2013).

[3]For example, Brewer and Ehrenberg (1996), Dearden et al. (2002), Monks (2000), Behrman et al. (1996), Loury and Garman (1995), and James et al. (1989).

Table 1: Wage Premium (%) to University Selectivity by Estimation Method and Country

| Country | Study | LS | ME | FE | SR | AM | IV | RDD |
|---|---|---|---|---|---|---|---|---|
| China | Hartoga et al. (2010) | 28[g] | | | | | | |
| US | Brewer et al. (1999) | 20[sd] | | | | | | |
| US | Long (2010) | 5[sd] | | | | | | |
| US | Andrews et al. (2016) | 10–21[g] | | | | | | |
| US | Black et al. (2005) | 11–13[g] | | | | | | |
| US | Black and Smith (2004) | 11–16[g] | 15–25[g] | | | | | |
| US | Brand and Halaby (2006) | 19[g] | 14[g,IN] | | | | | |
| US | Long (2008) | 2[sd]; 22.4[g] | 13.4[g] | | -13[sd,IN] | | 26[sd,IN] | |
| US | Dale and Krueger (2002) | 8[h] | | | -2[h,IN] | -0.1[h,IN] | | |
| US | Dale and Krueger (2014) | 6[h,LR] | | | -2[LR,IN] | | | |
| US | Behrman et al. (1996) | 21.2[g] | | 19.1[g] | | | | |
| US | Hoekstra (2009) | | | | | | | 18[g] |
| Chile | Bordon and Braga (2014) | 11.5[sd]; 26[g] | | | | | | 6.5–8[g] |
| Colombia | Saavedra (2009) | | | | | | | 20[g] |
| Israel | Lang and Siniver (2011) | 9–21[g] | | | | | | |
| UK | Chevalier and Conlon (2003) | 6[g] | | | | | | |
| UK | Hussain et al. (2009) | 6[sd] | | | | | | |
| UK | Walker and Zhu (2017) | 10[g,LR] | | | | | | |
| UK | Britton et al. (2016) | 6[g,LR] | | | | | | |
| UK | Broecke (2012) | 9.7[sd]; 20[g] | | | 5–7[sd]; 16–22[g] | 4–6[sd] | | |
| Italy | Triventi and Trivellato (2012) | | 2[g,IN] | | | | | |
| Italy | Brunello and Cappellari (2008) | 7.4[g] | | | | | | |
| Norway | Borgen (2014) | 2[sd] | | 1.5[sd] | 2[sd] | | 5[sd,LR] | |
| Sweden | Lindahl and Regnér (2005) | 7[g] | | 4[g] | | | | |
| Australia | Birch et al. (2009) | 0.6–1.6[g] | | | | | | |

LS= Least Squares estimator with an identification strategy that relies on a rich set of covariate.
ME= Average Treatment on the Treated effect estimated by the nearest neigbor (NNM) or the propensity score (PSM) matching.
FE= Sibling or Twin Fixed Effects estimator.
SR= Dale and Krueger (2002) Self-revelation Model
AM= Dale and Krueger (2002) Applicant-Matched Model
IV= Instrumental Variable estimator using the exclusion restriction as proposed in Long (2008) and inspired by Card (1999).
RDD= Regression Discontinuity Design that is originally proposed by Thistlewaite and Campbell (1960). Lee and Lemieux (2010) provide a comprehensive review of the technique.
[IN]= Indicates that the estimate is statistically insignificant and not different from zero.
[LR]= Indicates estimates on earnings that are recorded more than 5 years after graduation
[sd]= Indicates that the premium estimate corresponds to per one standard deviation of the composite quality index/university characteristic used in the respective paper.
[h]= Indicates that the premium estimate corresponds to per one hundred points of the average SAT/GMAT/entry exam score used in the respective paper.
[g]= Indicates that the premium estimate corresponds to an elite/top university grouping with respect to the prestige/reputation/ranking that is determined by the category/type of the institution or by the top quartile of the selectivity measure used in the respective paper.

(2004) is to allow for a non-restrictive functional form, in contrast with the conventional linear least squares estimator (LS). The weakness of both methods is its reliance on the available information in the data and the "selection of observables" assumption. Among the studies that follow these two estimation methods, Brewer et al. (1999) and Lang and Siniver (2011) differ slightly in their approach. Brewer et al. (1999) use a model based correction to address the self-selection issue but they find a significant premium of 20% that remains unaffected by the correction. Lang and Siniver (2011) investigates the earnings of an elite university's graduates, and compares those to a non-elite college that they claim is similar in many aspects except for the reputation in the labor market. This may be viewed as a match based on university characteristics, rather than individuals' characteristics.

The sibling/twin fixed effects (FE) estimator, first introduced in Behrman et al. (1996), is a within-household estimator and the FE purges all the genetically identical and unobserved characteristics that the siblings share. This approach implicitly assumes that the university choices are random within family (see Harmon et al., 2003, for the same discussion on returns to years of education). However, Goodman et al. (2015) shows that this is clearly not the case. This is why both Behrman et al. (1996) and Borgen (2014) find that the FE estimates are statistically not different from the LS estimates.

The SR model, and more so the AM model, are viewed as a more rigorous way of controlling to the unobserved motivation, aspiration and ambition of the students, which is assumed to be revealed by the average selectivity of the universities where each individual applied but was not admitted. The AM model is more specific in that it explicitly matches the students that have been accepted and rejected by the same universities and compares their earnings. Using U.S. data these two methods delivers low and insignificant estimates of the premium (Dale and Krueger, 2002, 2014; Long, 2008), but using U.K. data it delivers a substantial and statistically significant premium.

There have been two attempts in the literature so far with the use of instrumental variables. Long (2008) estimates a high but insignificant premium for the U.S. universities, but Borgen (2014) finds a 5% increase on the earnings of Norwegian graduates as the university selectivity measure increases by a standard deviation. As I discuss in section 3.3, the reason for the high estimates is related to the local nature of the IV estimator that is guided by the source of exogeneity that the exclusion restriction highlights.

Finally, the RDD method is viewed as the "closest cousin" (Lee and Lemieux, 2010) of experimental designs. However, external validity is often absent and can lead to high premium estimates of 18–20% as in Hoekstra (2009) and Saavedra (2009). This may be due to the fact that the studies can only use enrollees of one institution (the most selective) in the country as the treatment group, which leads one to question the external validity of the estimates. In contrast, Bordon and Braga (2014) are able to use information from all universities and estimate a wage premium to university selectivity of 6.5–8%.

**HE System Context** The tabulation of estimates by country of data and estimation method highlights the fact that the context of higher education (HE) systems being analyzed is an important dimension of variation in premium estimates. The main differences are in the following characteristics: (i) the source of funding (public funds vs. private tuition), (ii) the instruments they use to select students for admission (high school grades, standardized university entry exams, and other information related to non-analytical skills of the students), (iii) the dominance of the private HE sector in the country and its relative reputation in comparison to the public sector, (iv) and the variation in tuition fees. The relative variance of tuition fees, rather than the relative mean, may be used as a measure for the degree of differentiation among the HE systems, while relying on the assumption that price (tuition) reflects the quality of the product (human capital transmitted by universities).

Based on the above features, HE systems may be divided into three groups. The first group includes the severely selective systems with high variation in tuition fees, a private and reputable university sector, and use as admission criteria the scores on a standardized entry exam. Examples include China, Chile, the United States, the United Kingdom (OECD, 2007; Hartoga et al., 2010). The HE system of Israel is also similar in admission procedures as the countries above, except that the public institutions are the most selective ones that also ask for lower tuition fees. Therefore, characteristics of the tuition fees distribution is not informative of the selectivity differentiation in this case, as otherwise discussed above.

The second group includes those HE systems that are in the most part, if not fully, supported by public funds (i.e. tuition-free) where obtaining HE is considered a human right, not a privilege. Examples are the European Nordic countries like Sweden, Norway and

Finland (OECD, 2016, pg. 238). The universities in this group admit students based on their high school GPA. These countries have a very small private sector that is less selective than the public institutions. The third group includes the HE systems that borrow similarities from the above two HE system types and have moderately differentiated universities. Examples are Italy and Australia (Triventi and Trivellato, 2012; Birch et al., 2009; Hoxby, 1997). In these countries the universities' main source of funds is the government, the private sector is small (e.g.: 10% of enrollment in Italy), the tuition fees are moderate and have a low variation, and students are admitted based on their high school grades.

The Canadian HE system shares similar features with the third group. Higher education in Canada[4] is a provincial (not federal) jurisdiction, which led to the potentially most decentralized system in the developed world. Although there is no unique national system of education (each of the provinces have their own), the university sector features a homogeneous and uniform structure with a relatively low variation in public resources endowed to each institution. Tuition fees are lower than those of the universities in the first group, and have a much lower variation. There are relatively few universities in number[5] and students are admitted to universities based on their high school grade point average. In line with the above discussion, the estimates found in the literature that I summarize in Table 1 are higher for the countries with a strongly differentiated university system such as the U.S., Chile, Colombia, Israel and the U.K. The selectivity premium is smallest for Australia, and moderate for Italy, Norway and Sweden. In this paper I show that the selectivity premium for Canada is closest in magnitude to the universities in the latter group.

# 3 Methodology

## 3.1 Regression Model

In this section, I describe the estimation procedures that are used in this paper. I estimate equation (1), that defines individual $i$'s earnings in logarithmic form, $log(w)$, as a function of

[4] For a historical overview of the evolution of the Canadian higher education system see Jones (2014).

[5] There are 77 public and 13 private universities across Canada. Because of the limited data available on university characteristics, I have 46 universities in my sample, all of which are public. In the US, this number is much higher; for instance Black and Smith (2006) report 398 colleges in their sample only.

university quality, $S_j$, the set of individual, family and institution characteristics, $X_{ij}$, and an idiosyncratic error term, $u_{ij}$. The main interest of the paper lies in estimating parameter $\beta_1$ in equation (1).

$$log(w)_{ij} = \beta_0 + \beta_1 S_j + \beta_2 X_{ij} + u_{ij} \tag{1}$$

I use three different estimation strategies in pursuit of correcting for the sorting of students into better universities based on individual characteristics and ability. I use least squares and matching methods to correct this endogeneity based on a very rich set of student characteristics that are available in the data. I also use an instrumental variable approach to control for the potential presence of unobserved (and thus omitted) variables that could guide sorting. I describe below the later two approaches.

## 3.2 Matching Estimator

Black and Smith (2004) point out that matching methods outperform least squares by allowing for a non-linear selection on observable characteristics and thus do not impose a functional form on the type of dependency between the outcome and the covariates. They also show that the two methods yield different results. King and Nielsen (2016) show that matching on the propensity score lags behind by other types of matching. Following these suggestions in the literature I employ a multivariate, nearest neighbor matching (NNM) estimator. The underlying assumption of matching methods is that, conditional on the pre-treatment variables, assignment to treatment is independent of the outcome of interest. The matching estimator requires that the treatment variable be of a binary nature, indicating the treatment and the comparison group. The method finds matches with similar traits and attributes in the comparison group for each treated individual in the sample. In this paper, I select the closest matches (or neighbors) by minimizing the Euclidian distance between the characteristics of the treated individual and of similar neighbors in the comparison group. Hence, the treatment and the comparison group must contain sufficient overlap in the values of the pre-treatment variables for matching to be feasible. I allow matching with replacement, i.e. each individual in the comparison group may be used more than once as a match for different treated individuals. This reduces the estimation bias by yielding better matches. Once the closest matches have been selected for each treated individual, the counterfactual is estimated as the mean of their outcomes (wages). The mean of the differences in wages of the treated individuals and the corresponding counterfactual is the estimates of the Average Treatment on the Treated (ATT) parameter.

Empirical papers using matching methods report almost always the ATT estimate, assuming that in theory it should be equal to the Average Treatment on the Control (ATC) parameter. In a randomized control trial experiment these two parameters are equivalent. In this paper I compare the estimates of ATT and ATC (Average Treatment on the Control) to provide important insights with respect to the balancing property of the sample.

## 3.3 Instrumental Variables

In order to take into account any remaining sorting, not accounted by the available information in the data, I use an instrumental variable technique. Leveraging the available information the YITS, I construct two instruments (exclusion restriction). The first is similar to the one used in Long (2008). The author proposed the use of the average selectivity of universities within the residential area of the student during high school. Since students tend to attend a university not far away from their parents' home, and possibly in commuting distance, the distance from home may play a crucial role when choosing a university. The reasons are obviously related to travel and living costs given that Canada is a large and dispersed country. The geographic region in my case is the province where the student studied the last two years of high school. The instrument in this case exploits the exogenous variation in sorting that originates due to the university selectivity differences across provinces.

This instrument highlights the wage premium earned by the students who attended a better university because of the availability of better universities in proximity. The comparison group in this case are the students who went to a less selective university because of a high proportion of low ranked universities in their province, which lowers their likelihood of attending a highly ranked university. For these students (stayers) the return on their investment for a BA degree is higher than for the students that have to invest on moving and living costs away from their parents' household in addition to tuition fees. However, if these students (movers) are willing to undergo these investments in order to attend a more selective university, then they must be expecting that the returns would compensate for this investment.

With this in mind, the second instrument I construct exploits the difference in average university selectivity between the province of the students' high school and that of the university. The variable takes a positive value when students moved (in order to attend university) to a province with higher average university selectivity than their high school province, a zero value (by construction) when students choose to study university in the same province as their high school, and a negative value for those that attended a university in a province with lower average university selectivity than their high school province. This instrument leads to a LATE that weights more the students that reallocate to a province with a higher mean-selectivity and also attend selective university. The aim is to estimate a LATE that is representative of both students that attend a selective university just because many are availability in proximity, and of those students that are willing to invest more (living cost in addition to tuition) and reallocate a different province.

# 4 Data

## 4.1 Individual characteristics

The main data set used in this paper is the older cohort of the Youth in Transition Survey (YITS). Individuals born between 1979–1981 were surveyed every two years starting in 2000. They were then followed for eight years until the last interview, which was conducted in 2008.

The final sample excludes individuals with disabilities (4.3%), high school dropouts (4.81%), high school graduates (24.54%), and those whose educational status is unknown (2.89%). In the last cycle of the YITS, among the individuals who had completed some post-secondary education (PSE), 47% reported having completed a PSE program which is less than a Bachelor's degree (college degree, vocational, trades etc.)[6] and 15% a program that is higher than a Bachelor's degree (First Professional Degree, Master's, Ph.D., or post-graduate certifi-

[6]Canada has the highest rate of tertiary education qualification in the adult population (51% in 2011, (OECD, 2013)), and this is mostly due to the community college education system available in Canada.

cate/diploma). Since their wage structure is different from what it is for a regular Bachelor's degree graduate, these two groups are excluded from the final sample. For the main analysis, I focus on respondents who have attained a Bachelor's degree. In the end, the sample of Bachelor's degree graduates who were working full-time is composed of 1,476 individuals who were 26–28 years old by December 2007.

The outcome of interest is the hourly wage rate of the respondents 4–6 years after graduating from a university program and are reported as of December 2007. The unusual quality and information in the data allow for a large set of covariates to be included as controls in the regression analysis. These include:

(i) measures of academic aptitude: the high school grades (High school GPA), and the grades at the end of the first year in the university (GPA in First Year University). These are reported in the following categories: 90-100%, 80-89%, 70-79%, 60-69%, and 50-59% or lower intervals. In the regression model I use dummy variables indicating the two highest categories [7] ;

(ii) measures/proxies for inter-personal skills: an indicator variable if the individual ever volunteered (Volunteering), a measure in hour intervals per week[8] dedicated to participation in school clubs, teams, or school organizations (Participation in School Clubs) and non-school clubs[9] (Participation in Non-school Clubs), a self-reported 1–5 Likert scale rating of own writing and speaking skills in explaining ideas to others, speaking to an audience and participating in discussions (Communication Skills), a self-reported 1–5 Likert scale rating of own ability to identify problems and possible causes, to plan strategies, to solve problems or think of new ways to solve problems (Problem Solving Skills), a self-reported 1–5 Likert scale rating of own ability to use formulas to solve problems, interpret graphs or tables, using math to figure out practical things in everyday life (Math Ability). The last three variables enter the regression equation as a

[7]Since high school grades are self-reported, there is always the risk that they may be overstated. However, in the YITS the students were asked to report a grade interval. This procedure significantly reduces the risk of measurement error.

[8]Categories are from 0–6, where 0 indicates no participation, 1 indicates less than 1 hour per week, 2 indicates between 1–3 hours per week, 3 indicates between 4–7 hours per week, 4 indicates between 8–14 hours and 5 indicates more than 15 hours per week.

[9]These include non-school clubs, teams, volunteer work or other organizations, for example community sports, music lessons or youth groups not organized through school.

dummy indicator for reporting very good to excellent rating;

(iii) demographic characteristics: age of the individuals (Age), a gender indicator for females, race/community indicators (Black, Aboriginal, Visible minority), an indicator for command in both English and French (English and French speaker), an indicator for disability that does not impede any aspect of working (Disability indicator), a single marital status indicator, an indicator for rural residence (Rural), and province of residence (British Columbia, Quebec, Manitoba and Saskatchewan, Alberta, or Atlantic Provinces, Ontario is the reference category);

(iv) family characteristics: an indicator of at least one parent with some PSE qualification (Parents PSE), an indicator if student thought that getting PSE was important for their parents (PSE important for parent), and the number of older siblings, which is used as a proxy for the distribution of income within a family;

(v) high school characteristics: a private high school indicator (High school private), an indicator for consuming alcohol more than once a week during high school (Alcohol Consumption Indicator) which is used as a proxy for misbehavior, and a peer-effect variable reported by the student that their closest high school friends planned to continue their education into PSE (Friends plan PSE) (see e.g. Christofides et al., 2015);

(vi) university-related variables: duration in months of undergraduate studies as shown to be important by Black and Smith (2004) and Black et al. (2005), dummies for three major groups of fields of study: Social Sciences, Humanities and Arts, Business Administration and Commerce, and STEM (Sciences, Technology, Engineering and Mathematics);

(vii) job characteristics: experience in months following graduation (Experience), a set of dummies indicating three occupation groups: Management, Business, Finance and Administration, Natural and Applied Sciences, and Social Sciences, Education and Governmental Service. The omitted category is the fourth group: Art, Culture, Recreation and Sports; Sales, Service and other.

Table A.1 contains the summary statistics for the aforementioned individual characteristics. On average they are 27 years old. The estimation sample is composed of 60% females, 7.5% visible minority, 1% black and aboriginal respondents, 5% report to have a disability that does not restrict their full working potential, more than 80% report having older siblings and 50% of them have at least one parent with PSE education, and 83% report that attaining higher education is important for their parents. In the sample 7% of the students attended a private high school, more than 70% have a high school GPA higher than 80%, and more than half rate themselves as very good or excellent in communication, problem-solving and math skills. About 64% report having volunteered at least once, 23% to have consumed alcohol at least once a week while in high school, and about half report to have friends with PSE aspirations. On average individuals spend 1–3 hours/week engaged in school clubs, and about 4–7 hours/week in non-school clubs. With respect to their PSE experience, about 30% report a GPA in their first year university to be above 80%. On average the undergraduate studies take 13 semesters to complete. The average hourly wage is 24.6 dollars per hour. The respondents report on average 42 months (or 3.5 years) of post-graduation work experience. The sample has an almost uniform distribution across the occupation and field of study groups.

## 4.2 University characteristics and the quality measures

A remaining challenge in the literature of university selectivity premium is to properly measure selectivity. Several university characteristics are used individually or all together in a regression equation to capture the quality of educational institutions. Examples include faculty-student ratio, faculty salaries, enrollment, retention rate, publications per researcher, and applicants per place (Behrman et al., 1996; Dearden et al., 2002; Betts et al., 2013). Noticing a high correlation between these characteristics, some papers either use a single, most important variable (Dale and Krueger, 2002; Long, 2010) or use data reduction techniques to combine several characteristics in one comprehensive index (Black and Smith, 2004; Long, 2008). Other papers use a published quality ranking (Brewer et al., 1999; Monks, 2000). In fact Black and Smith (2006) presents evidence in favor of using several proxies rather than a single one, and if one should be used they argue that the average university entry exam score is the most reliable measure of selectivity in the U.S. I use three different selectivity measures: (1) the average GPA of entering cohort (GPA, (2) the *Maclean's* magazine Overall Reputation Ranking (RR) and (3) the Composite Index (CI) that I construct using several university characteristics.

I link the YITS to university characteristics from publicly available data. From the university ranking issue of *Maclean's* magazine published in November 2004 (Magazine, 2004) I retrieve the following variables: the Overall Reputation Ranking, the average high school grade point average of the entering cohort (Entering Cohort GPA), the percentage of full-time faculty with a Ph.D. degree (Faculty with Ph.D.), the number of students and full-time professors per 1000 who have won national awards in the past five years (Student Awards and Faculty Awards, respectively), and the number of Social Sciences and Humanities Research Council and Canada Council research grants per 100 full-time faculty members (Number Research Grants). From the 2002 CAUT Almanac (CAUT, 2002) I retrieve the ratio of full-time tenured faculty to the number of students enrolled (Faculty-Student Ratio).

First, in Canada the main admission criteria is the high school GPA. Therefore, I use the average high school grade point average as one of the selectivity measures. Second, the Reputation Ranking, is constructed by the *Maclean's* magazine. They conduct a survey on high school counselors, university officials, CEOs and corporate recruiters across Canada and ask them to rank universities based on three attributes: best quality, most innovative, and leaders of tomorrow. Then, *Maclean's* calculates a best overall Reputation Ranking as a simple average of the rankings on the three attributes. Therefore, this measure is based mainly on the experience that the individuals surveyed have had with graduates from the different universities and subjective perceptions of the relative quality of the human capital transmitted.

Third, the composite index is viewed as an objective measure of university selectivity. Following the literature (Black and Smith, 2004, 2006; Long, 2008), I construct this index by using a dimension reduction technique, in this case the Principal Component Analysis (PCA), to combine a set of university characteristics into a single index. PCA yields linear orthogonal combinations of the variables by assigning weights to each. These weights are determined by the solution of an optimization problem that maximizes the extent to which the index accounts for the correlation (and not the covariance since the inputs are in different units) between university characteristics. I use the first principal component of the orthogonal transformation as the Composite Index. This is an efficient and commonly used method in the literature to combine many university characteristics into one index without worrying about multicollinearity when, otherwise, they would be used jointly as covariates in a regression equation. Table A.2 provides a list of the university characteristics that I use to build the Composite Index and their descriptive statistics.

The *Maclean's* magazine classifies Canadian universities into three types: (1) Medical/Doctoral Universities, which are the universities that have medical schools and a broad range of PhD programs and research; (2) Comprehensive Universities, which are the universities that have a significant degree of research activity and a wide range of programs at the undergraduate levels, including professional degrees; and (3) Primarily Undergraduate Universities, which includes the universities that are largely focused on undergraduate education, with relatively fewer graduate programs and graduate students.

In order to define the university selectivity groups, I use the above classification of university together with the selectivity measured described above. Table 2 summarizes the definition of selectivity groups, which are build separately for each of the three measures (GPA, RR and CI). Group A includes the Medical/Doctoral universities that also rank within the top (forth) quartile of the selectivity measure's distribution. Group B includes the Comprehensive universities that rank in the top half of the selectivity measure's distribution and the Medical/Doctoral universities that rank within the third quartile. Group C includes all remaining universities.[10]

Table 2: Explaining the ranking group clasification

| | | University Type | | |
|---|---|---|---|---|
| | | Medical/Doctoral | Comprehensive | Primarily Undergraduate |
| Selectivity Measure | Fourth Quartile | Group A | | |
| | Third Quartile | | Group B | |
| | Bottom 50% | | | Group C |

Students in the YITS graduated from 93 different universities. Even though the university data from the magazines were available for only 49 universities, I could match 79% of the students in the YITS to their universities in the external data. The high match rate is because the magazine reports on the biggest universities in the country, that are also the ones that agree to participate in the ranking. The number of observations (students per university)

[10]In favor of space the list of universities in each group is omitted from the paper and is readily available from the author.

were very low for three universities which I exclude from the final sample.

# 5 Results

## 5.1 Matching estimates

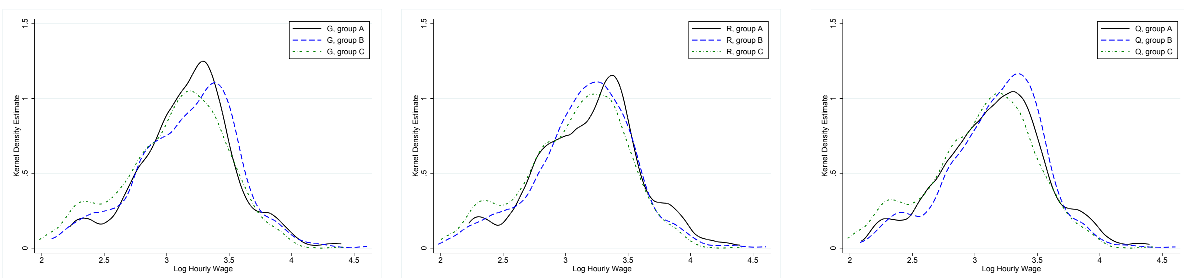
In this section I discuss the multivariate matching estimates of the university selectivity premium on wages along with the least squares estimates, a benchmark in the literature (see Table 1). Table 3 displays the wage premium estimates. The matching estimator requires that the treatment variable takes the form of a bivariate zero-one indicator. Therefore, I construct the treatment indicator to be a dummy variable. The first treatment variable takes the value one if the university that the student graduated from belongs to Group A, and zero if it belongs to Group B or C (columns (1)–(3) in Table 3). The second treatment variable takes a value one if the university belongs to Group A or B, and zero if it belongs to Group C (columns (4)–(6) in Table 3).

Table 3: Least squares and matching estimates of wage premium

| | Group A vs. B & C | | | Group A & B vs. C | | |
|---|---|---|---|---|---|---|
| | GPA | RR | CI | GPA | RR | CI |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Least Squares Estimates | | | | | | |
| coef. | 0.031 | 0.004 | 0.005 | 0.030 | 0.013 | 0.027 |
| s.e. | 0.026 | 0.030 | 0.029 | 0.026 | 0.025 | 0.024 |
| R-sq | 0.287 | 0.286 | 0.286 | 0.287 | 0.286 | 0.287 |
| Matching Estimates, ATT | | | | | | |
| coef. | 0.032 | 0.041 | 0.016 | 0.067*** | 0.059*** | 0.072*** |
| s.e. | 0.028 | 0.032 | 0.029 | 0.023 | 0.023 | 0.024 |
| Matching Estimates, ATC | | | | | | |
| coef. | 0.094*** | 0.056* | 0.069* | 0.091*** | 0.113*** | 0.099*** |
| s.e. | 0.033 | 0.030 | 0.027 | 0.024 | 0.024 | 0.023 |
| N | 1476 | 1476 | 1476 | 1476 | 1476 | 1476 |

*Note:* Treatment and comparison groups are indicated at the top of the columns. Standard error estimates are weighted and robust. The list of covariates not shown in the table is as listed and described in section 4.1. The treatment variables are defined as shown is Table 2 using the university type and the respective selectivity measure: Entering Cohort GPA (GPA), Reputation Ranking (RR), and the Composite Index (CI). ATT= Average Treatment on the Treated. ATC= Average Treatment on the Control. coef.= Coefficient estimate. s.e.= Standard error estimate. R-sq= R-squared. N= Sample size. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$.

The least squares (LS) regressions include all covariates described in Table A.1 and the quadratic of experience. The LS estimates in Table 3 indicate a 1–3% premium on the hourly wage rate, but it is statistically insignificant in all cases. I plot in Figure 1 the wage distribution by the selectivity group (A, B and C) of the graduates for each of the three selectivity measures (GPA, RR, CI). There is no obvious differences in the graduates earnings distributions. Group C university graduates' distribution is centered more to the left than the other two, but the difference is negligible. Therefore, the LS estimates are in line with the similarities of the distributions shown in Figure 1.




(a) GPA (b) RR (c) CI

Figure 1: Wage distribution by university selectivity group

The LS estimator weights each observation equally, in contrast with the matching estimator and the IV estimator (discussed in the next subsection) which weight the observations non-equally. In addition, the matching estimator also does not impose the linear functional form restriction as LS estimator does. Differently from the LS estimates, the average treatment effects on the treated (ATT) estimates in Table 3 are higher in magnitude, and statistically significant in the last three columns. Graduating from a Group A or B university earns their alumni a 6–7% premium on the hourly wage rate, when compared to Group C university graduates. The estimate is consistent across the three different selectivity measures (GPA, RR, CI).

I report the ATC estimates in the third panel of Table 3, as a diagnostic check for the balancing property on the treatment and comparison groups. In the ideal setting of experimental data, the two groups would be well balanced (i.e. very similar in terms of observable characteristics) and the estimates of ATT and ATC would be identical as a result. This is because it would not matter if we decide to find a match for each treated individual, or for each untreated individual since due to the composition of the two groups would lead one to find identical matches in both cases. As seen from Table 3, the estimates of ATT and ATC, eve though not identical, they yield similar results, which in most cases are statistically not different (except for column (1) and (2)). The more commonly used balance test for matching estimates (Smith and Todd, 2005) is left out of the paper in favor of space, but the tables are available from the author.

## 5.2 Instrumental variable estimates

Following the description of the estimation approach in Section 3.3, in this subsection I discuss the instrumental variable estimates that are shown in Table 4. In the third panel of the table I show the unconditional correlation coefficients between the endogenous variable and the instruments, which are higher than those reported in Long (2008) and Borgen (2014), vary between 0.13–0.63. The first stage regressions indicate that the instruments are both relevant, except for column (2). I report the Hansen J-test of overidentifying restrictions (Hansen, 1982). In all cases the test indicates that the instruments are valid and not correlated with the error term.

The IV estimates indicate a 8% premium on the hourly wage for those that graduated from a Group A university in comparison to other universities. The premium is 14.8% when the treatment is defined as graduating from a Group A or B university versus a Group C university. In both cases the treatment groups are defined using the Entering Cohort GPA as the selectivity measure. For the other cases when the treatment definition is based on the RR and the CI (columns (2), (3),(5) and (6)), the premium varies between 6–14.4% but is statistically insignificant. Finally, with respect to the estimate in colum (4), even though the IV estimate is almost twice the ATT estimate of Table 3, they are statistically not different.

## 5.3 Robustness analysis

Analyzing the distribution of the predicted probability of attending a better university as defined by each treatment variable, commonly referred to as the propensity score (PS), may allow us to look into another aspect of the student composition of the treatment and comparison groups. The PS is simply the predicted outcome from a logit equation of the treatment indicator variable on all covariates that I use in the multivariate matching to construct Table

Table 4: Instrumental variable estimates of wage premium

| | Group A vs. B & C | | | Group A & B vs. C | | |
|---|---|---|---|---|---|---|
| | GPA | RR | CI | GPA | RR | CI |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| IV-GMM Second Stage | | | | | | |
| coef. | 0.077** | 0.144 | 0.127 | 0.148* | 0.060 | 0.107 |
| s.e. | 0.035 | 0.180 | 0.141 | 0.078 | 0.038 | 0.079 |
| R-sq | 0.276 | 0.263 | 0.264 | 0.264 | 0.275 | 0.272 |
| First Stage | | | | | | |
| ΔAverage Selectivity Measure | 0.607*** | 0.068 | 0.347** | 0.218 | 0.917*** | 0.552** |
| s.e. | 0.038 | 0.091 | 0.166 | 0.135 | 0.087 | 0.210 |
| Average Selectivity Measure | 0.269*** | 0.540* | 0.512** | 0.804*** | 0.737*** | 0.728*** |
| s.e. | 0.089 | 0.280 | 0.248 | 0.157 | 0.179 | 0.161 |
| R-sq | 0.442 | 0.216 | 0.297 | 0.426 | 0.507 | 0.450 |
| Unconditional correlation coefficient between the endogenous variable and the instruments: | | | | | | |
| ΔAverage Selectivity Measure | 0.63 | 0.13 | 0.41 | 0.43 | 0.58 | 0.44 |
| Average Selectivity Measure | 0.38 | 0.34 | 0.34 | 0.56 | 0.34 | 0.52 |
| Overidentification Test p-value | 0.798 | 0.199 | 0.367 | 0.231 | 0.864 | 0.384 |
| N | 1476 | 1476 | 1476 | 1476 | 1476 | 1476 |

*Note:* Treatment and comparison groups are indicated at the top of the columns. Standard error estimates are weighted and robust. The list of covariates not shown in the table is as listed and described in section 4.1. The treatment variables are defined as shown is Table 2 using the university type and the respective selectivity measure: Entering Cohort GPA (GPA), Reputation Ranking (RR), and the Composite Index (CI). ATT= Average Treatment on the Treated. ATC= Average Treatment on the Control. coef.= Coefficient estimate. s.e.= Standard error estimate. R-sq= R-squared. N= Sample size. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$.

3 estimates. The kernel density estimates are displayed in Figure A.1 and A.2. In the former the PS distribution for the treatment group is an almost symmetric bell shaped, whereas the distribution of the comparison group is right-skewed. In the latter the PS distribution for the treatment group is left-skewed and the one for the comparison group is right skewed. One may be concerned on whether the NNM estimates presented in Table 3 are driven by and reflect mostly the students with high PS. As a robustness check, I trimmed the tails of the PS distribution and kept the students whose PS score fell within the inter-quartile range (IQR). The empirical distributions are shown in Figures A.3 and A.4.

Using the IQR sample I reproduce the estimation results of Tables 3 and 4, and display the estimates in Tables A.3 and A.4. These estimates are very similar to the previous estimates that I retrieve using the full sample: LS estimates hover between 1–4% and are statistically insignificant; the NNM estimates of ATT effect are between 6–8% and statistically significant at conventional levels; ATC estimates are statistically equal to ATT estimates; the IV estimates provide a statistically premium of 8–13%.

## 5.4 Selection after enrollment

The paper so far has discussed and addressed the issue of non-random selection of students in universities of different selectivity at the enrollment stage. What happens in the stages that follow after enrollment could also potentially lead to non-random selection.

Dropping out of university is likely to be systematically different across universities of different quality. Also, some of the students may take longer to graduate than others, and once they do, they may enter the labor market or pursue graduate studies. Those who enter the labor market may be working full time or part time. Any or all of these choices are also sources of sample selection which could potentially distort the estimation results if they are systematically correlated with university quality. In order to test the presence of this type of non-random selection, Table A.5 displays the regression results of several outcomes on the dummy indicators for the highest- and middle-third university groups, leaving lowest-third universities as the omitted category. These outcomes are: the probability of dropping out of university, time in months it took to complete the program and graduate, the probability of pursuing graduate school versus entering the labor market, the probability of being employed at the time of the survey, and the probability of being self-employed.

According to the grouping based on the Reputation ranking, group A university graduates are 6% less likely to drop out. Nevertheless, the outcome is not consistent across the other two types of groupings. The non-random selection due to dropping out behavior could be an issue with important consequences when working with data from countries where university retention and completion rates are low (e.g. Symonds et al. (2011) report a 56% retention rate in the U.S.). Canada's universities have high retention rates of above 80%. When both RR and the CI is used to form selectivity groups, graduates of group A universities are 5–6% more likely to continue further their education with a graduate degree, and are about 4% more likely to find employment after graduation (the sample excludes the students enrolled in a graduate degree program). The former may lead to an estimate of the wage premium that is downward biased (since the better students of the top universities are out of the labor market), and the latter may lead to an overestimate (since group A graduates would be more present than group B& C graduates). As long as these two effects translate into an equal proportion of bias in the wage premium, these may potentially cancel out. Lastly, from the third panel of Table A.5, graduates of group A or B universities take 0.05 months longer to graduate. This is a sufficiently small magnitude that is unlikely to affect the wage premium estimates.

# 6 Conclusion

Abundant evidence in the literature documents a positive and high wage premium to university selectivity. The majority of the studies use U.S. data or data from countries with a higher education system that are strongly differentiated. In this paper I survey the literature on wage premium of university selectivity and argue that not only the identification strategy and the method of estimation matter, but the characteristics of the different higher education systems also matter. I observe that the wage premium is lower in countries like Canada where the higher education is funded heavily through public funds and the universities are moderately differentiated in terms of selectivity.

In addition, I use the Youth in Transition Survey, a unique and previously unexploited data to estimate the wage premium to university selectivity in Canada. This is the first paper which attempts to tackle the issue of selection on unobservable characteristics in the context of Canada. Using a matching estimator I estimate a 7% premium on the hourly wages of graduates to selective Canadian universities. The instrumental variable estimates, that are free of bias from unobserved variables, indicate a 14.8% premium. The two estimates are statistically not different.

A question of immediate interest that has not been rigorously investigated in the literature is whether the positive premium of university selectivity changes in later career when the wage profile reaches equilibrium. Another question is related to separating between the potential channels that lead wages of observationally similar individuals to be 7–14.8% higher: Is it the quantity/quality of human capital transfered from the more selective universities or is it a beneficiary side effect of being in a network of better peers? Or is it simply signaling? Few recent efforts on early career wages provide support for the signaling scenario (Lang and Siniver, 2011; Bordon and Braga, 2014; Macleod et al., 2017) without unambiguously separating or ruling out the other potential mechanisms. These research questions are subject of future research upon availability of new data that permit such analysis.

# References


Rodney J. Andrews, Jing Li, and Michael F. Lovenheim. Quantile treatment effects of college quality in earnings. *The Journal of Human Resources*, 51(1):200–238, 2016.

Jere R. Behrman, Mark R. Rosenzweig, and Paul Taubman. College choice and wages: Estimates using data on female twins. *Review of Economics and Statistics*, 78(4):672–685, 1996.

Julian Betts, Christopher Ferrall, and Ross Finnie. The role of university characteristics in determining post-graduation outcomes: Panel evidence from three Canadian cohorts. *Canadian Public Policy*, 39(s1):81–106, 2013.

Elisa Rose Birch, Ian Li, and Paul W. Miller. The influences of institution attended and field of study on graduates' starting salaries. *The Australian Economic Review*, 42(1):42–63, 2009.

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.

Paola Bordon and Breno Braga. Employer learning, statistical discrimination and university prestige. *Manuscript (April 2014)*, 2014.

Nicolai T. Borgen. College quality and hourly wages: Evidence from the self-revelation model, sibling models and instrumental variables. *Social Science Research*, 48:121—-134, 2014.

Jennie E. Brand and Charles N. Halaby. Regression and matching estimates of the effects of elite college attendance on educational and career achievement. *Social Science Research*, 35:749–770, 2006.

Dominic Brewer and Ronald Ehrenberg. *Research in Labor Economics*, volume 15, chapter Does it pay to attend an elite private college? Evidence from the senior class of 1980, pages 239–271. JAI Press, Greenwich, CT, 1996.

Dominic J. Brewer, Eric R. Eide, and Ronald G. Ehrenberg. Does it pay to attend an elite private college? *Journal of Human Resources*, pages 104–123, 1999.

Jack Britton, Lorraine Dearden, Neil Shephard, and Anna Vignoles. How English domiciled graduate earnings vary with gender, institution attended, subject and socio-economic background. April 2016. URL `http://www.ifs.org.uk/uploads/publications/wps/wp201606.pdf`.

Stijn Broecke. University selectivity and earnings: Evidence from UK data on applications and admissions to university. *Economics of Education Review*, 31:96–107, 2012.

Giorgio Brunello and Lorenzo Cappellari. The labour market effects of Alma Mater: Evidence from Italy. *Economics of Education Review*, 27:564–574, 2008.

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.

CAUT. CAUT almanac of post-secondary education, 2002. URL `https://www.caut.ca/docs/almanac/almanac-2002.pdf?sfvrsn=0`.

Weiwei Chen, Wayne A. Groveb, and Andrew Hussey. The payoff to school selectivity: An application of Dale and Krueger's method to MBA programs. *Economics Letters*, 116: 247–249, 2012.

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 Christofides, Michael Hoy, Joniada Milla, and Thanasis Stengos. Grades, aspriations and post-secondary education outcomes. *Canadian Journal of Higher Education*, 45(1): 48–82, 2015.

Sarah R. Cohodes and Joshua S. Goodman. Merit aid, college quality, and college completion: Massachusetts' Adams Scholarship as an in-kind subsidy. *American Economic Journal: Applied Economics*, 6(4):251–85, October 2014. doi: 10.1257/app.6.4.251. URL `http://www.aeaweb.org/articles?id=10.1257/app.6.4.251`.

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), 2002.

Stacy B. Dale and Alan B. Krueger. Estimating the effects of college characteristics over the career using administrative earnings data. *The Journal of Human Resources*, 49(2): 324–358, 2014.

Lorraine Dearden, Javier Ferri, and Costas Meghir. The effect of school quality on educational attainment and wages. *The Review of Economics and Statistics*, 84(1):1–20, 2002.

Joshua Goodman, Michael Hurwitz, Jonathan Smith, and Julia Fox. The relationship between siblings' college choices: Evidence from one million SAT-taking families. *Economics of Education Review*, 48:75–85, 2015. ISSN 0272-7757. doi: http://dx.doi.org/10.1016/j.econedurev.2015.05.006. URL `http://www.sciencedirect.com/science/article/pii/S0272775715000722`.

Wayne A. Grove and Andrew Hussey. Returns to field of study versus school quality: MBA selection on observed and unobserved heterogeneity. *Economic Inquiry*, 49(3):730–749, 2011. ISSN 1465-7295. doi: 10.1111/j.1465-7295.2010.00292.x. URL `http://dx.doi.org/10.1111/j.1465-7295.2010.00292.x`.

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054, 1982.

Eric A. Hanushek, Victor Lavy, and Kohtaro Hitomi. Do students care about school quality? Determinants of dropout behavior in developing countries. *Journal of Human Capital*, 2 (1):69–105, 2008.

Colm Harmon, Hessel Oosterbeek, and Ian Walker. The returns to education: Microeconomics. *Journal of Economic Surveys*, 17(2):115–155, 2003.

Joop Hartoga, Yuze Sunb, and Xiaohao Ding. University rank and bachelor's labour market positions in China. *Economics of Education Review*, 29:971–979, 2010.

James J. Heckman and Richard Robb. Alternative methods for evaluating the impact of interventions. *Journal of Econometrics*, 3:239–267, 1985.

Scott Heil, Liza Reisel, and Paul Attewell. College selectivity and degree completion. *American Educational Research Journal*, 51(5):913–935, 2014.

Mark Hoekstra. The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics*, 91(4):717–724, 2009.

Caroline Hoxby. how the changing market structure of u.s. higher education explains college tuition, 1997.

I. Hussain, Sandra McNally, and Shqiponja Telhaj. University quality and graduate wages in the UK, 2009.

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.

Glen A. Jones. An introduction to higher education in Canada. In K.M. Joshi and Saee Paivandi, editors, *Higher education across nations*. Delhi: B.R. Publishing, 2014.

Katja Maria Kaufmann, Matthias Messnerz, and Alex Solis. Returns to elite higher education in the marriage market: Evidence from Chile. *Unpublished Mimeo*, 2013.

Katja Maria Kaufmann, Matthias Messner, and Alex Solis. The intergenerational effects of elite higher education: Evidence from Chile. *Manuscript (March 2014)*, 2014.

Gary King and Richard Nielsen. Why propensity scores should not be used for matching. *Working Paper*, December 2016. URL `https://gking.harvard.edu/files/gking/files/psnot.pdf`.

Kevin Lang and Erez Siniver. Why is an elite undergraduate education valuable? Evidence from Israel. *Labour Economics*, 18(6):767–777, 2011.

David S. Lee and Thomas Lemieux. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(June):281–355, 2010. doi: http:www.aeaweb.org/articles.php?doi=10.1257/jel.48.2.281.

Lena Lindahl and Høakan Regnér. College choice and subsequent earnings: Results using Swedish sibling data. *Scandinavian Journal of Economics*, 107(3):437–457, 2005.

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.

Linda Datcher Loury and David Garman. College selectivity and earnings. *Journal of Labour Economics*, 13(2):289–308, 1995.

Prashant Loyalka, Yingquan Song, and Jianguo Wei. The effects of attending selective college tiers in china. *Social Science Research*, 41(2):287 – 305, 2012. ISSN 0049-089X. doi: http://dx.doi.org/10.1016/j.ssresearch.2011.11.015. URL `http://www.sciencedirect.com/science/article/pii/S0049089X11002274`.

W. Bentley Macleod, Evan Riehl, Juan E. Saavedra, and Miguel Urquiola. The big sort: College reputation and labor market outcomes. *American Economic Journal: Applied Economics*, 9(3):223–261, 2017.

Maclean's Magazine. 2004 University Rankings. *Maclean's*, 117(46):26–72, 2004.

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.

OECD. Oecd thematic review of tertiary education, 2007. URL `http://www.oecd.org/edu/skills-beyond-school/39355618.pdf`.

OECD. Education at a glance, 2013. URL `http://www.oecd.org/edu/Canada_EAG2013%20Country%20Note.pdf`.

OECD. Education at a glance 2016: OECD indicators, 2016.

Anna Rosefsky Saavedra and Juan Esteban Saavedra. Do colleges cultivate critical thinking, problem solving, writing and interpersonal skills? *Economics of Education Review*, 30: 1516–1526, 2011.

Juan E. Saavedra. The learning and early labor market effects of college quality: a regression discontinuity analysis. *Unpublished manuscript, Universidad de los Andes, School of Government*, 2009.

Jeffrey A. Smith and Petra E. Todd. Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*, 125:305—-353, 2005.

Alex Solis. Does higher education cause political participation? Evidence from a regression discontinuity design, 2013. URL `http://uu.diva-portal.org/smash/get/diva2:638723/FULLTEXT01.pdf`. Retrieved: 23-08-2017.

William C. Symonds, Robert Schwartz, and Ronald F. Ferguson. Pathways to prosperity: Meeting the challenge of preparing young Americans for the 21st century, 2011. URL `http://nrs.harvard.edu/urn-3:HUL.InstRepos:4740480`. Retrieved: 05-02-2017.

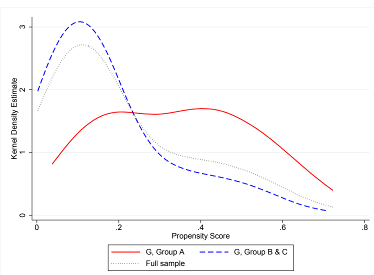
D. Thistlewaite and D. Campbell. Regression-discontinuity analysis: An alternative to the ex-post facto experiment. *Journal of Educational Psychology*, 51(6):309–317, 1960.

Moris Triventi and Paolo Trivellato. Does graduating from a private university make a difference? Evidence from Italy. *European Journal of Education*, 47(2):260–276, 2012.

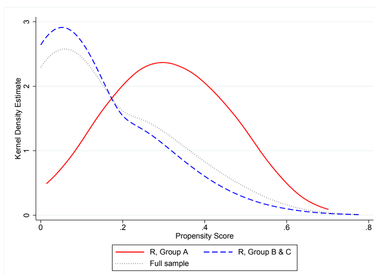
Ian Walker and Yu Zhu. University selectivity and the graduate wage premium: Evidence from the UK, 2017.

# Appendix

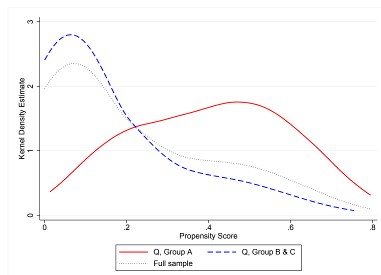
Figure A.1: Propensity score distribution of Group A vs. B & C




(a) GPA

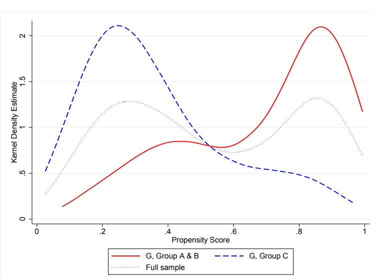



(b) RR

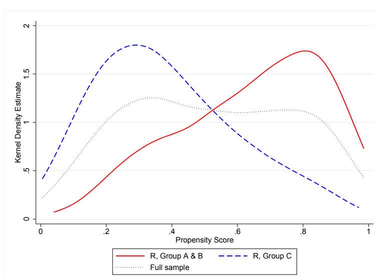



(c) CI

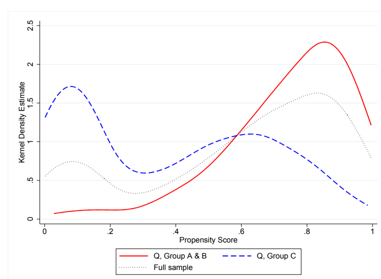
Figure A.2: Propensity score distribution of Group A & B vs. C




(a) GPA

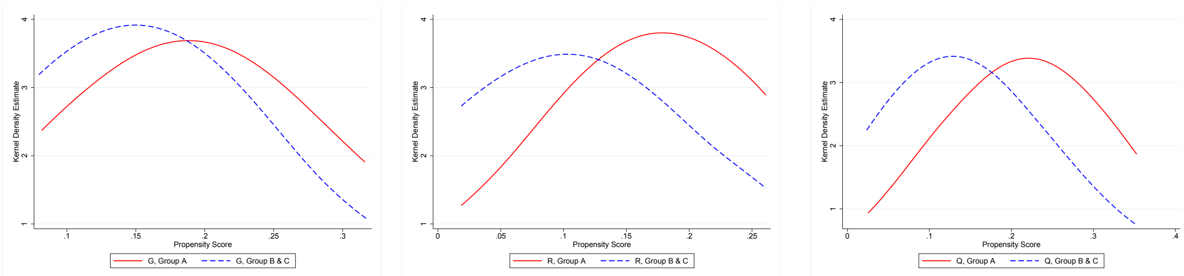



(b) RR



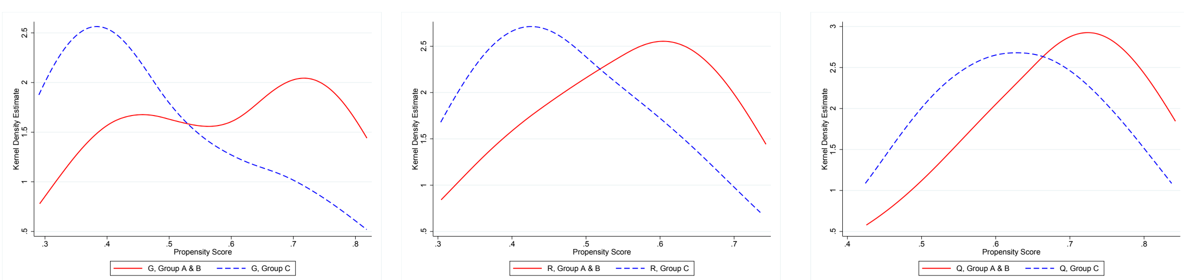

(c) CI

Figure A.3: Propensity score distribution of Group A vs. B & C of the IQR




(a) GPA (b) RR (c) CI

Figure A.4: Propensity score distribution of Group A & B vs. C of the IQR




(a) GPA (b) RR (c) CI

Table A.1: Summary Statistics: Mean (Standard Deviation)

| Variable | Mean | Std.Dev. |
|---|---|---|
| **Demographic variables**: | | |
| Age | 26.986 | 0.798 |
| Female indicator | 0.605 | 0.489 |
| Visible minority indicator | 0.075 | 0.263 |
| Aboriginal indicator | 0.010 | 0.097 |
| Black indicator | 0.011 | 0.104 |
| Both English & French Speaker indicator | 0.453 | 0.498 |
| Married indicator | 0.506 | 0.499 |
| Disability indicator | 0.045 | 0.208 |
| Number of Older siblings | 0.831 | 1.072 |
| Parents PSE indicator | 0.512 | 0.491 |
| PSE very important for parents | 0.831 | 0.374 |
| Rural Residence indicator | 0.147 | 0.354 |
| Residence Province: British Columbia | 0.071 | 0.257 |
| Residence Province: Quebec | 0.141 | 0.348 |
| Residence Province: Manitoba, Saskatchewan | 0.125 | 0.331 |
| Residence Province: Alberta | 0.135 | 0.342 |
| Residence Province: Atlantic Provinces | 0.189 | 0.392 |
| **High school-related variables**: | | |
| Private High School Indicator | 0.071 | 0.256 |
| High School GPA 90-100% | 0.182 | 0.386 |
| High School GPA 80-89% | 0.523 | 0.499 |
| Communication Skills | 0.671 | 0.470 |
| Problem Solving Skills | 0.780 | 0.415 |
| Math Ability | 0.530 | 0.499 |
| Volunteer indicator | 0.642 | 0.479 |
| Participation in School Clubs | 2.484 | 1.547 |
| Participation in Non-School Clubs | 3.183 | 1.500 |
| Alcohol Consumption Indicator | 0.233 | 0.423 |
| All Friends Plan PSE | 0.414 | 0.492 |
| Most Friends Plan PSE | 0.462 | 0.498 |
| **Undergraduate studies-related variables**: | | |
| GPA in 1st Year University 90-100% | 0.045 | 0.204 |
| GPA in 1st Year University 80-89% | 0.259 | 0.433 |
| Bachelor's Duration in Months | 52.698 | 5.140 |
| Major: Social Sciences, Humanities and Arts | 0.182 | 0.386 |
| Major: Bussiness Administration, Commerce | 0.190 | 0.392 |
| Major: STEM | 0.280 | 0.449 |
| **Labor market-related variables**: | | |
| Hourly Wage | 24.569 | 10.030 |
| Experience in Months | 41.871 | 23.676 |
| Occ.: Management, Business, Finance and Administration | 0.269 | 0.444 |
| Occ.: Natural and Applied Sciences | 0.268 | 0.443 |
| Occ.: Social Sciences, Education and Governmental Services | 0.263 | 0.440 |
| N | 1476 | |

Table A.2: Summary statistics and the weighting scheme of university characteristics

| Variable | Composite Index Weight | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|---|
| Entering Cohort GPA | 0.431 | 81.617 | 3.467 | 75 | 89 |
| Faculty-Student Ratio | 0.083 | 0.198 | 0.043 | 0.100 | 0.280 |
| Faculty with PhDs | 0.369 | 88.078 | 10.866 | 38.2 | 98.400 |
| Student Awards | 0.494 | 3.563 | 2.307 | 0.200 | 9.500 |
| Faculty Awards | 0.462 | 3.534 | 3.021 | 0 | 10.700 |
| SSHR Number | 0.462 | 16.615 | 11.051 | 1.500 | 47.730 |

Table A.3: Least squares and matching estimates of wage premium, IQR sample

| | Group A vs. B & C | | | Group A & B vs. C | | |
|---|---|---|---|---|---|---|
| | GPA | RR | CI | GPA | RR | CI |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Least Squares Estimator | | | | | | |
| coef. | 0.041 | 0.006 | 0.038 | 0.034 | 0.038 | 0.038 |
| s.e. | 0.037 | 0.046 | 0.030 | 0.029 | 0.030 | 0.028 |
| R-sq | 0.312 | 0.282 | 0.274 | 0.303 | 0.289 | 0.287 |
| Matching Estimates, ATT | | | | | | |
| coef. | 0.070 | 0.023 | 0.033 | 0.059* | 0.078** | 0.058* |
| s.e. | 0.045 | 0.053 | 0.046 | 0.033 | 0.033 | 0.034 |
| Matching Estimates, ATC | | | | | | |
| coef. | 0.102*** | 0.063 | 0.061 | 0.106*** | 0.078** | 0.069* |
| s.e. | 0.040 | 0.051 | 0.051 | 0.033 | 0.032 | 0.035 |
| N | 738 | 738 | 738 | 738 | 738 | 738 |

*Note:* Treatment and comparison groups are indicated at the top of the columns. Standard error estimates are weighted and robust. The list of covariates not shown in the table is as listed and described in section 4.1. The treatment variables are defined as shown is Table 2 using the university type and the respective selectivity measure: Entering Cohort GPA (GPA), Reputation Ranking (RR), and the Composite Index (CI). IQR sample= The sample used for these results is composed of individuals with a propensity score (PS) estimate that falls within the Inter-quartile range of the PS distribution. ATT= Average Treatment on the Treated. ATC= Average Treatment on the Control. coef.= Coefficient estimate. s.e.= Standard error estimate. R-sq= R-squared. N= Sample size. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$.

Table A.4: Instrumental variable estimates of wage premium, IQR sample

| | Group A vs. B & C | | | Group A & B vs. C | | |
|---|---|---|---|---|---|---|
| | GPA | RR | CI | GPA | RR | CI |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| IV-GMM Second Stage | | | | | | |
| coef. | 0.089* | -0.386 | 0.021 | 0.131* | 0.080* | 0.107 |
| s.e. | 0.048 | 0.344 | 0.140 | 0.076 | 0.045 | 0.089 |
| R-sq | 0.300 | 0.260 | 0.287 | 0.277 | 0.291 | 0.273 |
| N | 738 | 738 | 738 | 738 | 738 | 738 |
| First Stage | | | | | | |
| ΔAverage Selectivity Measure | 0.608*** | 0.080 | 0.388** | 0.236* | 0.885*** | 0.664*** |
| s.e. | 0.042 | 0.097 | 0.161 | 0.136 | 0.109 | 0.222 |
| Average Selectivity Measure | 0.322*** | 0.254* | 0.493** | 0.936*** | 0.885*** | 0.837*** |
| s.e. | 0.108 | 0.135 | 0.236 | 0.204 | 0.218 | 0.192 |
| R-sq | 0.390 | 0.135 | 0.209 | 0.311 | 0.415 | 0.393 |
| Overidentification Test | 0.720 | 0.357 | 0.012 | 0.363 | 0.743 | 0.242 |
| N | 738 | 738 | 738 | 738 | 738 | 738 |

*Note:* Treatment and comparison groups are indicated at the top of the columns. Standard error estimates are weighted and robust. The list of covariates not shown in the table is as listed and described in section 4.1. The treatment variables are defined as shown is Table 2 using the university type and the respective selectivity measure: Entering Cohort GPA (GPA), Reputation Ranking (RR), and the Composite Index (CI). IQR sample= The sample used for these results is composed of individuals with a propensity score (PS) estimate that falls within the Inter-quartile range of the PS distribution. ATT= Average Treatment on the Treated. ATC= Average Treatment on the Control. coef.= Coefficient estimate. s.e.= Standard error estimate. R-sq= R-squared. N= Sample size. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$.

Table A.5: Testing for sample selction after enrollment

| | Group A vs. B & C | | | Group A & B vs. C | | |
|---|---|---|---|---|---|---|
| | GPA | RR | CI | GPA | RR | CI |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Pr(Drop out of university program) | | | | | | |
| coef. | -0.011 | -0.060*** | -0.019 | -0.013 | 0.003 | -0.013 |
| s.e. | 0.016 | 0.020 | 0.016 | 0.014 | 0.013 | 0.014 |
| (Pseudo) R-sq | 0.229 | 0.235 | 0.229 | 0.229 | 0.228 | 0.229 |
| N | 2230 | 2230 | 2230 | 2230 | 2230 | 2230 |
| Pr(Graduate degree) | | | | | | |
| coef. | 0.005 | 0.051** | 0.060*** | 0.015 | -0.007 | 0.015 |
| s.e. | 0.021 | 0.023 | 0.021 | 0.020 | 0.018 | 0.021 |
| (Pseudo) R-sq | 0.045 | 0.047 | 0.049 | 0.045 | 0.045 | 0.045 |
| N | 2089 | 2089 | 2089 | 2089 | 2089 | 2089 |
| Bachelor's duration in months | | | | | | |
| coef. | -0.004 | -0.032 | -0.030 | 0.003 | 0.049*** | 0.012 |
| s.e. | 0.017 | 0.021 | 0.019 | 0.017 | 0.016 | 0.018 |
| R-sq | 0.036 | 0.036 | 0.036 | 0.036 | 0.038 | 0.036 |
| N | 2089 | 2089 | 2089 | 2089 | 2089 | 2089 |
| Pr(Employment) | | | | | | |
| coef. | 0.019 | 0.041*** | 0.043*** | 0.016 | 0.013 | 0.017 |
| s.e. | 0.014 | 0.015 | 0.015 | 0.015 | 0.014 | 0.016 |
| (Pseudo) R-sq | 0.103 | 0.110 | 0.112 | 0.103 | 0.103 | 0.103 |
| N | 1679 | 1679 | 1679 | 1679 | 1679 | 1679 |
| Pr(Self-employed) | | | | | | |
| coef. | 0.007 | 0.002 | 0.004 | 0.002 | 0.011 | 0.005 |
| s.e. | 0.011 | 0.013 | 0.012 | 0.012 | 0.012 | 0.012 |
| (Pseudo) R-sq | 0.176 | 0.175 | 0.175 | 0.175 | 0.177 | 0.175 |
| N | 1505 | 1505 | 1505 | 1505 | 1505 | 1505 |

*Note:* Treatment and comparison groups are indicated at the top of the columns. Standard error estimates are weighted and robust. The list of covariates not shown in the table is as listed and described in section 4.1. The treatment variables are defined as shown is Table 2 using the university type and the respective selectivity measure: Entering Cohort GPA (GPA), Reputation Ranking (RR), and the Composite Index (CI). coef.= Coefficient estimate. s.e.= Standard error estimate. R-sq= R-squared. N= Sample size. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.