Finish It and It's Free: An Evaluation of College Graduation Subsidies

Matthew D. Webb*

December 15, 2014

Abstract

Despite the rapid increase in the returns to higher education witnessed in the labor market over the past few decades, there has also been a marked increase in the share of individuals who dropout of college or university. To boost student persistence in higher education, several Canadian provincial governments introduced a set of reforms that were designed as subsidies for college graduation. In addition, these policies were designed to discourage internal migration following graduation. Using data from both administrative tax records as well as longitudinal surveys, I analyze the effectiveness of these policies. The main findings are that the programs had no effect on internal migration, but did significantly reduce college dropout rates.

Preliminary - Please do not cite.

The current version of this paper can be found on my website: `https://sites.google.com/site/matthewdwebb/`

*Department of Economics, University of Calgary, Calgary, Alberta, Canada, T2N 1N4. Email: mwebb@ucalgary.ca. I thank my supervisors, James MacKinnon and Steven Lehrer, for their continued support. I am grateful to Michele Campolieti, Marco Cozzi, and Allan Gregory for thoughtful evaluations on a prior draft. I am grateful to participants at the 40th and 42nd ACEA conferences, the 47th CEA meetings, and the 84th SEA Conference. I appreciate the suggestions from Emilia Simeono, Fraser Summerfield, and Julian Hsu. Thank you to David R. Johnson's and René Morrisette for providing me with data and code. Part of this analysis was conducted at the Prairie Regional RDC which is part of the Canadian Research Data Centre Network (CRDCN), I am thankful for the assistance of Bin Hu and Rebecca Williams. Paul Roberts' help at Statistics Canada is also greatly appreciated. I gratefully acknowledge financial support from the Social Sciences and Humanities Research Council of Canada.

1 Introduction

In the past several decades there has been a large increase in college and university enrollment. There has also been a large increase in the college dropout rate (Turner, 2004). Furthermore, completion rates have stagnated among recent cohorts, as discussed in Oreopoulos and Petronijevic (2013). Currently, the six-year graduation rate in four-year colleges in the US is 58% and the seven-year completion rate at a large sample of Canadian universities is 84%.[1] With dropout rates being as large as they are, many have started to ask what can be done to increase college persistence.[2] This paper examines a set of tax policies which included strong incentives for college graduation, offering tax credits totalling between $15,000 and $25,000. Four Canadian provinces have offered these credits to recent graduates in an effort to curb outflows of graduates to other provinces. This paper analyzes the impact that these programs had on college enrollment, college dropout, and migration decisions.

These programs have similar aims to the various Merit Scholarship programs offered by several American state governments.[3] Both the Merit Scholarships and the retention credits attempt to raise the average level of educational attainment within a jurisdiction. The merit programs and the retention programs both tie educational funding to a geographic location but the merit programs do not restrict where recipients can reside post-graduation. The Merit Scholarships offer incentives to enroll in college in a specified location, while the retention credits offer incentives to graduate from college and to reside in a specified location after graduation. Additionally, the Merit Scholarships offer financial assistance during the years of college enrollment, while the retention credits offer assistance after graduation. While the Merit Scholarships have received considerable attention, there has been very little analysis of the Graduate Retention Credits.[4] This paper provides the first analysis of causal outcomes of these programs.

Concern has been given to the local stock of human capital in recent years, as an important factor for regional economic growth and development. In particular, several researchers have studied the effectiveness of various policies at increasing the local stock.[5] The lock stock of graduates has proved difficult to increase as higher levels of education also increase an

[1]The American figure is from National Center for Education Statistics (2012) and the Canadian figure is the author's calculation using data from Maclean's Magazine (2014).

[2]See for example, Bailey and Dynarski (2011), Attewell, Heil and Reisel (2011), and Turner (2004) among many others.

[3]Fifteen states now offer some form of merit scholarship, and the effectiveness of these various programs has been studied by Fitzpatrick and Jones (2012).

[4]Essaji and Neill (2010) detail the costs and program features of these credits but not the consequences of these programs.

[5]See for instance, Groen (2011), and Winters (2014).

individuals' likelihood to emigrate within country (Malamud and Wozniak, 2012). Other research has suggested that Merit Scholarships do not increase the retention of college-bound students after graduation(Sjoquist and Winters, 2013b). In part, the retention programs were introduced to discourage internal migration since historically many graduates have moved provinces after they graduate.[6] This paper additionally analyzes whether the programs changed migration patterns of recent graduates, and finds the patterns unchanged.

Using both administrative and survey data the programs are found not to have an impact on migration decisions. This perhaps explains why one of the provinces, Nova Scotia, recently cancelled its graduate retention program.[7] This paper also introduces a new proxy for college graduation into the Longitudinal Administrative Databank. The analysis shows that the programs did have a significant impact on reducing college dropout rates, by 4.1% among individuals 18-23. Interestingly, there was no estimated increase in the likelihood of enrolling in college. These results are opposite of what is typically found when tuition is reduced, where enrollment increases, but completion rates stagnate.[8]

The remainder of this paper is as follows, section 2 places the programs in the context of the literature. Section 3 gives a detailed overview of the various graduate retention programs under study. Section 4 describes the linear Difference-in-Differences (DiD) estimation strategy and the relevant treatment and comparison groups. As inference with DiD is difficult, especially with few treatment and comparison groups, the methodology for inference is discussed in section 4.3. Section 5 describes the administrative and survey data sources and the outcomes of interest. Results are discussed in section 6, which present evidence from both data sources that the programs had a significant impact on decreasing college dropout rates, but did not decrease out of province migration. Section 7 concludes.

2 Context of the Retention Programs

This paper investigates how individuals respond to subsidies for college graduation. Starting in 2005, several Canadian provinces offered 'Graduate Retention Credits', in an effort to discourage recent graduates from moving out of province. The design of these programs offered large tax credits which were conditional on graduation. The programs are explained in greater detail in section 3. In general, the programs offered between $15,000 and $25,000 in tax credits to recent university graduates. The credits refunded between 50% and 97% of

[6] The Saskatchewan government mentions that there "has been significant leakage of post-secondary graduates outside the province"(The Saskatchewan Labour Market Commission's, 2009).

[7] See `http://www.novascotia.ca/finance/en/home/taxation/tax101/personalincometax/grr.aspx` for more details.

[8] Cohodes and Goodman (2014) find completion rates decreased among those treated by the Massachusetts Merit Scholarship.

average four-year total tuition in these provinces. This paper examines the impact that these programs had on a host of education and migration decisions. It finds that the programs decreased dropout rates by 4.1%.

The witnessed increase in dropout rates is surprising given that the relative returns to education have increased over the same time period. Deming and Dynarski (2009) find that real wages for high school graduates fell by one-third from 1972 to 2005, while real wages for college graduates held steady. For men, the bachelor's degree wage premium was 22% in 1972 and it had increased to 60% by 2003. Findings such as these, lead Oreopoulos and Petronijevic (2013) to summarize that although the returns are heterogeneous, college is a worthwhile investment for both the average and the marginal student.

The increase in the dropout rate would not be as worrying if the benefits of going to college increased linearly in years of education. However, it has been known since Hungerford and Solon (1987) that there are "degree effects" or "sheepskin effects" conferred on those who graduate.[9] More recent literature has found that degree effects exist across the distribution of earnings, Oreopoulos and Petronijevic (2013), and that earnings for those with some-college are only slightly higher than the earnings of high school graduates. Within Canada, Ferrer and Riddell (2008) estimate sheepskin effects in wages for holders of bachelor's degrees on the order of 20% for women and 16% for men. Additionally, Riddell and Song (2011) find that re-employment prospects, following a job termination, are higher for college graduates. Finally, Jepsen, Troske and Coomes (2014) find both earnings premiums and higher levels of employment for community college graduates. The prevalence of sheepskin effects across a variety of countries and education levels suggests that individuals who dropout do so at significant private cost.[10] That the programs decreased dropout rates is surprising, given the large rewards in the labor market for graduating.

In addition to the private costs there are also social costs from dropping out. Schneider and Yin (2012) perform a 'back of the envelope' calculation and suggest large losses in income tax revenue from dropouts. There is also speculation that many new jobs will require higher levels of education suggesting diminished employment prospects for dropouts going forward. According to the United States Bureau of Labor Statistics, job growth in occupations requiring some post secondary education is expected to outpace job growth in occupations requiring a high school education or less over the coming decade (Bureau of

[9] The literature on sheepskin effects is vast and has found among other things that the effects are larger for females and minorities, Belman and Heywood (1991), Ferrer and Riddell (2008); and exist in several countries, Denny and Harmon (2001).

[10] However, it is possible that the dropout decision is welfare improving as simulations in Stinebrickner and Stinebrickner (2013) show that newly enrolled students who perform poorly learn that staying in school is not worthwhile.

Labour Statistics, U.S. Department of Labor., 2012).

This paper investigates whether the programs were able to retain recent graduates, as the programs were intended to do. There are many reasons why a provincial government would want to retain these individuals. The post-secondary education (PSE) spillovers literature presents evidence that higher levels of education benefits the community at large. When analyzing minimum schooling laws, Acemoglu and Angrist (2001) find evidence that there are small positive external returns to an extra year of education. Moretti (2004) uses the presence of a land grant university as an instrumental variable for college attendance to examine the impact of increasing the share of college graduates in a city. He shows that a percentage point increase in the presence of college graduates is associated with increased wages for others: specifically a 1.9% wage increase for high-school dropouts and a 1.6% increase for high school graduates. In a similar study, Shapiro (2006) examined American data at the metropolitan level and finds that a 10% increase in a city's concentration of college graduates was on average followed by a 0.8% increase in employment growth. Aghion et al. (2009) find that investment in four-year college educations has a positive effect on economic growth.

The programs are unique in that they place no restrictions on pre-college residency, or on where one studies, but restrict where one works after graduation. While most education funding for students has been offered with few geographical constraints, there have been a few exceptions. There are the aforementioned Merit Scholarships, and differential tuition levels for in-state and out-of-state students. Some jurisdictions, including Canadian provinces, have experimented with targeted retention/attraction programs to attract individuals in certain occupations, such as doctors and nurses.[11] Maryland has offered scholarships to residents which require an individual to work one year in the state for each year they receive a scholarship (Groen, 2011). In 2012, Kansas offered incentives such as student debt repayments and income tax waivers to attract individuals to rural Kansas. Unlike the graduate retention programs, to be eligible for the incentives individuals must prove they have resided outside of the state for at least the previous five years.[12] Finally, in 2007, Maine introduced a program generally similar to the retention credits studied here. This program repaid up to $5500 per year in student loans for bachelor degree holders from a Maine college.[13] The outcomes of either the program in Maine or Kansas have yet to been analyzed.

[11] See Reamy (1994) and Mullan (1999).

[12] For more information of the Kansas program see http://http://www.kansascommerce.com/index.aspx?nid=320

[13] For more information see http://www.opportunitymaine.org/opportunity-maine-program/frequently-asked-questions/.

Given the large value of these credits, it is interesting to compare the programs to other large scale education financing reforms. Gunnes, Kirkebøen and Rønning (2013) investigate an experiment in Norway where students were offered a financial incentive of $3000 (USD) if they graduated 'on time'. They find that the incentive reduced mean graduation delay by 0.23 semesters per year treated. Similarly, Garibaldi et al. (2012) find that at Bocconi University a 1000 € increase in tuition reduced the likelihood of late graduation by 5.2%, with no increase in the dropout rate. Arendt (2013) finds that a large increase in grants decreased dropout rates, but had no impact on completion rates after controlling for various student and parental characteristics. Dowd (2004) finds that an increased amount of subsidized loans in the first year of college enrollment increased persistence.

Past literature on financial aid suggests that is can be difficult to predict the impact of a reform as administrative details and program knowledge are important. Deming and Dynarski (2009) show that 'financial aid' is not a homogeneous good, and that paperwork matters. Programs with high administrative hurdles have smaller benefits. The retention credits are relatively easy to apply for, and in some provinces can be claimed as part of provincial tax returns. The salience of these programs, especially among high school students, is an open question. McGuigan, McNally and Wyness (2012) surveyed high school students in London, England after recent tuition reforms and found that roughly half of the students surveyed were unaware of key features of the reforms.

3 Background of the Graduate Retention Programs

The four Canadian provinces that have implemented graduate retention programs are Saskatchewan, Manitoba, New Brunswick and Nova Scotia.[14] Table 1 provides an overview of the various program attributes and how they differ across provinces.[15] Broadly speaking the programs are quite similar, and the amounts they offer to college and university graduates are of the same order of magnitude. All of the programs are income tax credits, though the characteristics of the credits differ across provinces. Only one of the credits is refundable, though most of the credits do roll-over.[16]

[14] See Essaji and Neill, 2010 for a summary of the various GRP programs.

[15] Quebec also operates a smaller wage subsidy program for people in remote, resource-rich regions who work in the resource industry, with eligibility contingent on holding a degree related to your current occupation. Given the specificity of this program it is ignored in the analysis. For more details about the Quebec program see http://www.revenuquebec.ca/en/citoyen/credits/credits/credits_reduisant/nouv_diplome/

[16] Specifically until 2012 Saskatchewan offered a refundable credit, which meant that if individuals earned insufficient income to claim the maximum annual amount, they would receive the difference in the form of a refund. In 2012, Saskatchewan allowed the credit to roll-over, but ended the refundability provision. All but Nova Scotia offers a roll-over provision, which allows the unused portion of the credit to accrue over time. This means that eventually an individual will be able to claim the maximum allowable credit.

Saskatchewan and Manitoba do not require a separate application for the graduate retention credits, which can be claimed on income tax returns. Nova Scotia and New Brunswick require separate applications to claim the credits. Three of the four programs determine the maximum credit based on the amount of tuition paid, while Nova Scotia offers a fixed amount to each recent graduate. The proportion of tuition refunded in each province varies, with up to 100% of tuition being refundable in Saskatchewan and 50% being refundable in New Brunswick. The maximum amount of the credits is the same in both of these provinces, but given the differing tuition refund percentages a student would have had to pay $40,000 in tuition to receive the maximum credit in New Brunswick, but only $20,000 in Saskatchewan. Finally, the total costs of each program are broadly similar in each of the provinces, ranging from $24-$35 million per year. These figures are estimated to increase over the coming years as many of these programs are new and there are not yet six or seven cohorts of graduates claiming these credits. In early 2014, the government of Nova Scotia eliminated its Graduate Retention Rebate. The government cites the failure of the programs to retain graduates as the reason for cancelling the programs.[17] This is perhaps not surprising as results in section 6 suggest that the various programs did not alter internal migration decisions.

Aside from the retention credits the provincial and federal governments invest heavily in both post-secondary institutes and students. Essaji and Neill (2010) provide a summary of the characteristics and costs of the various student funding programs currently in operation in Canada, and thoroughly summarize the graduate retention programs. In addition to the programs mentioned in that paper, several provinces have recently adjusted their approach to funding post-secondary education. For instance, Ontario recently introduced a tuition credit where students are eligible for a 30% rebate of their Ontario college or university tuition.[18] Additionally, proposed tuition increases in Quebec resulted in student strikes lasting for over six months. Further information about the details and costs of government funding for post secondary education can be found in Neill (2013).

The graduate retention programs are operationally quite different than the other means of government funding for post-secondary education. The most distinct feature is that they base eligibility on graduation. Another unique feature is that their benefits are provided solely after graduation. Ontario's rebates refund tuition in the year that it was paid. Similarly, the federal and provincial tuition tax credits and education amounts are claimed during the years enrolled in school. Both of these are non-refundable so they typically do not ease

[17] See `http://www.novascotia.ca/finance/en/home/taxation/tax101/personalincometax/grr.aspx` for more details.

[18] The grant tops out at $1680 for university tuition and $770 for college tuition, provided that their parents earn $160,000 or less per year, see `https://osap.gov.on.ca/OSAPPortal/en/PostsecondaryEducation/Tuition/index.htm` for more details.

a student's budget constraint until after graduation, as explained in Neill (2013). Student's budget constraints are a current concern in education policy, since it has been shown that increased financial aid improves college enrollment.(Sartarelli, 2011) The retention credits do nothing to ease this constraint while enrolled, but expand a student's post graduation budget. Offsetting the less than ideal timing of the benefits, is the fact that most student loans require repayment following graduation and the payouts from the programs coincide with the repayment schedule for student loans.[19]

The rebate schedules of these programs introduces an added complexity to calculating the costs of attending university. For example, if a student in Saskatchewan paid the average university tuition for four years starting in the fall of 2007, then her annual tuition would have been \$5015, \$5064, \$5173, and \$5431. Of the \$20,683 paid in tuition, they would be eligible for a \$20,000 rebate (or 96.69%), paid out in the following annual instalments: \$2000,\$2000,\$2000,\$2000,\$4000,\$4000,\$4000. Determining the net present cost of such a decision is perhaps not the easiest task for a student in their final year of high school. Moreover, in most cases individuals need to have a sizeable income to receive the full amount of the credit in the least amount of time.[20] Aspects of these programs are similar to the UK tuition reforms where the net present liability of tuition depends on labour market outcomes after leaving school.[21]

Another distinct feature of these programs is that they are place dependent. This is an unusual characteristic of education funding, but it does align the subsidy with the provinces' goals of having a well-educated labour force, rather than a large number of college and university students. While the programs are designed to retain an individual in a province, they are not targeted towards those on the margin of emigrating. As a result, the programs offer generous tuition rebates to all those who were never contemplating leaving the province. However, these rebates are conditional on graduation and thus may influence college graduation and dropout decisions. Having explained how these programs operate institutionally, it is worth considering how they may influence an individual's decisions theoretically.

4 Identification and Inference Strategy

I propose a simple model of how the introduction of graduate retention credits could impact an individual's decision making. I then describe how the impacts proposed in the model

[19]The analysis discussed in section 6 directly estimates the impact of the credits on interest paid for student loans.

[20]The roll-over provisions offered by many provinces mean that the entire value of the credit will eventually be received, but over a longer time horizon.

[21]In the UK, the full value of tuition is loaned to students while in school, and repayment is conditional on post schooling wages. (McGuigan, McNally and Wyness, 2012)

will be estimated using a linear difference-in-differences (DiD) equation. Finally, as inference with DiD can be difficult, the procedure for inference is discussed.

4.1 A Simple Model

Let us consider the situation of a student who has just finished high school. The individual wants to maximize her utility which is a function of both her discretionary income $\tilde{C}$ and quality of life (QoL). She is able to do so by choosing only two things: her level of education and the city in which she resides. Cities influence her quality of life, her earnings, and her taxes. Conceptually, we can imagine this individual solving the following problem:

$$\max_{\substack{educ_1 \in \{hs,c,u\} \\ educ_2 \in \{hs,c,C,u,U\} \\ city_{1,2,3} \in X}} U_1(\tilde{C}_1, QoL_1) + \beta U_2(\tilde{C}_2, QoL_2) + +\beta^2 U_3(\tilde{C}_3, QoL_3),$$

$$\begin{aligned}
&\text{such that} \\
&QoL_i = f(city_i), \\
&\tilde{C}_1 = Y_1(educ_1, city_1) + SL_1(educ_1) - \\
&\qquad taxes(city_1) - rent(city_1) - tuition(educ_1), \\
&\tilde{C}_2 = Y_2(educ_2, city_2) + SL_2(educ_2) - (1+r)SL_1(educ_1|educ_2) \\
&\qquad - taxes(city_2) - rent(city_2) - tuition(educ_2), \\
&\tilde{C}_3 = Y_3(educ_2, city_3) - (1+r) * SL_2(educ_2) \\
&\qquad - taxes(city_3, educ_2) - rent(city_3).
\end{aligned}$$

The problem is to maximize her utility across three periods; two periods in which she can possibly further her education and the final period. Utility in the second and third periods is discounted by an exogenously given discount factor, β. Simplifying, assume that she can only make the following educational choices: in the first period she can choose no further education and remain a high school graduate (hs), she can attend a community college (c), or she can attend university (u). Additionally, she can choose the city in which she will either work or attend school from the set of cities, X. In the second period, she can choose to either remain in school, or to dropout of school. If she chooses not to continue her education in the first period, by assumption, she remains at education level hs. If she attended college in the first period she can either enter the workforce with some college, c, or finish her program and graduate, C. Similarly, if she attended university in the first period she can either enter the workforce with some university, u, or finish her program and graduate, U.

In the first two periods the educational decision will affect consumption in several ways.

Obtaining additional education requires a good deal of time, and as such will negatively impact earnings, Y, in the first period. Additional education has a direct tuition cost, which must be paid in the current period. Student loans (SL) are available to help pay tuition and provide additional consumption in the first two periods if additional education is chosen, but they must be repaid with interest, r, in the third period. The individual is free to relocate to a new city after entering the workforce. Education is assumed to increase earnings in the third period. The graduate retention programs introduce an education argument to the tax function in the third period, as in certain cities the individual will face lower taxes, given earnings, than a similar person with a lower level of education on account of receiving a graduate retention tax credit. Rent is included in the model to reflect the fact that there are some non-traded goods that people must consume, the price of which is determined by the city of residence.

Reflecting on the impact of the graduate retention programs through the lens of this model reveals several dimensions in which the programs might have an impact. The most direct impact is that consumption in the third period is higher in a city where an individual is eligible for a credit, with all else being equal. Accordingly, if wages, taxes, and quality of life are all equal between two cities, then if one city offered a credit and the other did not we should expect the city with the credit to be the chosen residence. Moreover, the availability of a credit will increase overall consumption in the three periods for those who go to school, so we should expect that the introduction of a retention credit will increase educational attainment. This follows from the fact that when a location offers a credit, the credit increases overall three-period consumption for graduates. This increase in consumption makes obtaining further education relatively more attractive. A third possible impact may occur because the decrease in taxes from a credit coincides with student loan repayments. If there is heterogeneity in the amount of student loans outstanding, then its probable that those with larger amounts of debt would be more likely to locate in a city offering a credit. Similarly, those with student loans may be able to accelerate their debt repayment if they receive a credit.

The graduate retention programs have one additional channel of influence. If people have already decided to enroll in school in expectation of receiving a credit upon graduation then the costs of dropping out of school are more severe than they would have been otherwise. In the absence of a credit the potential costs of dropping out is the forgone higher wages that one might earn after graduating. The presence of a credit increases the cost of dropping out by disqualifying an individual from receiving the credit. One can alternatively think about the credits as a fee/rebate program, wherein payment is made up front, and a rebate is offered on graduation. If one expected to receive in the third period a refund proportional to her first period tuition in period three, then the decision to dropout in period two would retroactively increase the 'expected' cost of going to school in period one.

Any influence that a program may have will depend on the age at which an individual was when the program was announced. The model is written as though the program is in place in period zero. This will be the case for individuals who reach the end of high school in or after the year a program is announced. However, there are individuals for which the programs were announced when they were in period one or period two. For these individuals, the enrollment decision will have already been made, thus the programs are more likely to impact the migration and graduation decisions. The empirical analysis will account for these differences. The model above is framed in terms of an individual's decision, which matches well with the micro level data used for analysis. With these possible outcomes in mind, I now discuss the estimation strategy.

4.2 Estimating Equation

A linear difference-in-differences (DiD) approach will be used. Given that the impacts of the programs may depend on the age an individual was when the programs were announced, the analysis will be split into two parts. The first part will estimate the impact that the programs had on period one and two decisions, using data on individuals aged 18-23. The second part will estimate the impact the programs had on period three decisions, using data on individuals 23-28. Individuals who turned either 18 (or 23) in a province offering a retention credit will be regarded as 'treated'.

In the analysis, individuals in the provinces offering graduate retention programs listed in table 1 will be regarded as 'treated', while individuals in the other provinces will be regarded as the 'comparison' group. The linear DiD will then compare the changes in outcomes over time in provinces with retention credits to changes in outcomes over time in provinces without. Identification using DiD requires there to be common support between the treatment and comparison groups, and common trends prior to the introduction of the programs in both groups, see DiNardo and Lee (2011) for additional details.

The sample used for analysis is different for the early period and late period analysis, given differing concerns about common support. The early period analysis, which focuses on educational outcomes, uses individuals from all Canadian provinces.[22] The late period analysis, which focuses on migration decisions, will look at the impact of the programs in the Atlantic provinces. The Atlantic provinces offer a useful setting for conventional DiD

[22] As mentioned above, the majority of Canadians attend university within their own province. Accordingly, I assume that changes in the net cost of attending university in one province are unlikely to affect the decision of someone in another province to attend university.

analysis of migration. The Atlantic region is comprised of two provinces with retention programs: Nova Scotia and New Brunswick; and two provinces without: Prince Edward Island, and Newfoundland and Labrador. There is the additional benefit of interprovincial migration being nearly symmetric within the Atlantic Region.[23] Accordingly, the assumption of common trends in interprovincial migration is more realistic when restricting the sample to the Atlantic provinces.

The impacts of the credits on early period outcomes are estimated by the following equation:

$$\begin{aligned} Y_{ist} = c + \beta_{GRP} * I[\text{GRP Prov at } 18_{ist} * \text{age GRP announced}_{ist}] \\ + \text{TRENDS}_{st} + \text{PROV}_t + \text{AGE}_i + \text{BIRTH YEAR}_i + \epsilon_{ist} \end{aligned}$$

In the equation there is a set of province dummy variables, a set of birth year dummy variables, a set of province specific time trends, and a set of age dummy variables. Aside from age dummies there are no control variables, as the unconditional impact of the programs is of primary concern. The coefficient of interest is β_{GRP} the DiD coefficient. This coefficient will capture the marginal impact of being in a province with a retention credit, for those who were young enough to be eligible for a credit. The variable for this coefficient is an indicator variable for the interaction between living in a province offering a credit and being young when the credit was announced. This variable will be equal to one for those who turned 18: in Nova Scotia in $2006 - 2013$, in New Brunswick in $2005 - 2013$, and in Saskatchewan or Manitoba in $2007 - 2013$. The variable is set to zero otherwise. The set up for the late period estimating equation is nearly identical, but treatment status is defined by when an individual turned 23, and uses only individuals from the Atlantic provinces. This equation estimates a common treatment effect for the provinces offering graduate retention credits.

4.3 Inference Strategy

Inference with linear difference-in-differences can be a challenge. The prevalence of type I errors has been known since Donald and Lang (2007) and Bertrand, Duflo and Mullainathan (2004). The source of these problems is largely due to serial correlation in the error terms. These problems can be corrected for by clustering standard errors at the level of the policy change. Accordingly, the standard errors are calculated using a cluster robust variance estimator, clustered by province. However, the cluster robust variance estimator is unreliable when there are few clusters,[24] or when clusters are unbalanced.[25] In this analysis, clusters

[23] See Statistics Canada, CANSIM table 051-0019.
[24] See Cameron, Gelbach and Miller (2008).
[25] See MacKinnon and Webb (2014) and Carter, Schnepel and Steigerwald (2013).

are both severely unbalanced and there are only four or ten clusters in the dataset.

To deal with the small cluster problem, Cameron, Gelbach and Miller (2008) (CGM) propose a wild cluster bootstrap which they claim works with as few as five clusters. However, Webb (2014) shows that with few clusters, the 2-point wild cluster bootstrap is unreliable.[26] With few clusters, the CGM procedure calculates a point estimate for the p-value although the p-value is actually interval identified. Webb (2014) proposes the following 6-point distribution for use with few clusters:

$$v_g = -\sqrt{\frac{3}{2}}, -\sqrt{\frac{2}{2}}, -\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}}, \sqrt{\frac{2}{2}}, \sqrt{\frac{3}{2}} \qquad w.p. \qquad \frac{1}{6}.$$

To deal with unbalanced clusters, MacKinnon and Webb (2014) show that the wild cluster bootstrap works well. However, the simulations in that paper have a comparatively large number of clusters. Additionally, the simulations in Webb (2014) involve only balanced clusters. This leaves open the question of whether the 6-point distribution works well with few clusters. I preform a small Monte Carlo exercise to answer that question.

The Monte Carlo exercise uses the homoskedasticity set up from Cameron, Gelbach and Miller (2008). Across 25,000 replications, data are generated by the following equation:

$$y_{ig} = x_{ig} + x_g + \epsilon_{ig} + \epsilon_g$$

Where $x_{ig}, x_g, \epsilon_{ig}, \epsilon_g$ are all drawn from a $N(0,1)$ distribution. This set up imposes within cluster correlation in both the x-variable $(x_{ig} + x_g)$ and in ϵ $(\epsilon_{ig} + \epsilon_g)$. Two cases are considered, one to match the data in the early decision analysis, and one to match the late decision analysis. In both cases there are 1000 observations per replication. In the early decision analysis the data are divided among ten clusters, proportional to the populations of Canadian provinces. In the late decision analysis the observations are distributed proportionally to the populations of the four Atlantic provinces.

Within each replication the following regression is estimated:

$$y_{ig} = c + \beta_x x_{ig} + \mu_{ig}$$

The null hypothesis $\beta_x = 1$ is tested using several different procedures for inference. The 5% rejection rate is then compared across the various procedures. The first procedure is

[26] Simplifying, the CGM procedure generates bootstrap samples by transforming residuals, from a restricted regression, by a variable v_g, drawn from a bootstrap weight distribution. The same v_g is applied to all observations within a cluster. CGM recommend using the Rademacher distribution in which $v_g = \pm 1$ with probability 0.5.

to use cluster robust standard errors, and to assume that the t-statistic follows a $t(G-1)$ distribution, where G is the number of clusters.[27] The other procedures use the wild cluster bootstrap using either the Rademacher distribution of the 6-point distribution. Webb (2014) proposes enumerating the elements of the empirical distribution (as opposed to sampling the elements through bootstrapping) when there are few clusters. As a result, both the bootstrap and enumerated p-values are used for both distributions.

Results from the four cluster Monte Carlo design are in the top of Table 2. The rejection frequencies for the four cluster design show three things. The first is that with only four clusters the rejection frequency for the cluster robust procedure is far too large. Second, the p-value intervals from the enumerated 2-point distribution are quite wide. Third, the bootstrap procedures tend to under-reject. The second row of the table calculates a 5% critical value for each test. That it, is calculates the 5^{th} percentile of the 25,000 p-values for each of the procedures. These values show that the cluster robust p-values are far too low as 5% of the replications were estimated to be significant at the 99.95% level. Conversely, the 6-point bootstrap 5% critical value is 11.5% suggesting that far fewer than 5% of the replication were estimated to be significant at the 95% level. These simulations show that the cluster robust p-values significantly understate the p-value, while the bootstrap p-values will overstate the p-value.

Results from the ten cluster Monte Carlo design are in the bottom of Table 2. There are two many differences between the ten cluster and four cluster designs. The number of clusters is obviously greater, but now there are four 'treated' clusters, as opposed to only two treated clusters in the four cluster design. Results in MacKinnon and Webb (2014) show that the wild cluster bootstrap works well when the number of treated clusters is reasonably large. In the ten cluster case there is not the problem of under-rejection when using either the 2-point or the 6-point bootstrap. The cluster robust p-values still over reject too often, but the over-rejection is far less severe compared to the four cluster case. Finally, the 6-point wild cluster bootstrap is preferable to the 2-point wild cluster bootstrap, as the 2-point enumerated intervals are still somewhat wide with ten clusters.

The results from this Monte Carlo suggest that careful attention must be paid to inference. In the early decision analysis, when using data from all ten provinces, the wild cluster bootstrap will result in well-sized test. However, in the late decision analysis, when using data from only the four Atlantic provinces inference will be more difficult. In this case, the cluster robust p-values will be quite significantly understated, while the wild cluster bootstrap p-values will be over-stated. As a result, the cluster robust p-value can be thought of as a lower bound, and the wild cluster bootstrap p-value as an upper bound, with that

[27] This is the default procedure when the 'cluster' command is used in Stata.

interval containing the 'true' p-value.

5 Data

The primary dataset for this paper comes from Statistics Canada's Longitudinal Administrative Databank (LAD). The dataset uses administrative tax-file data on a 20% sample of Canadians. As the name suggests, the databank is longitudinal and contains tax information for each individual from $1982 - 2011$.[28] The dataset also records age, gender, and province of residence. The LAD allows for estimation of all of the outcomes of interest. The longitudinal nature of the data allows for a nearly perfect measure of migration.[29] Certain items in the tax code allow a researcher to infer university enrollment. The procedure for this estimation was developed by Morissette, Chan and Lu (forthcoming) and is explained in the Appendix. These same items can be used to infer whether someone graduated from university. Thus, the LAD can be used to estimate mobility, university enrollment, and educational attainment.

The sample used for the first part of the analysis is individuals aged $18 - 23$ in the years $2000 - 2011$, and the outcomes are summarized in Table 3. *Enrolled in University* is a binary indicator for whether an individual was enrolled in university during the current tax year. *Ever Enrolled in University* is a binary indicator for whether that individual was ever enrolled in university. *Enrolled in University for 2 years* is a binary indicator for whether the individual had been enrolled in university for at least two years, as measured by having been enrolled for 16 full-time equivalent months.[30] *University Graduate* is a binary indicator for whether the individual graduated from university, measured by the individual having been enrolled for 24 full-time equivalent months. *University Dropout* is a binary indicator for whether the individual had dropped out of university. The variable is set equal to one for individuals who have previously enrolled in university, have not graduated from university, and are not currently enrolled in university. *Student Loan Interest (to date)* measures the cumulative total of student loan interest that an individual has claimed on her tax returns. This variable is analyzed only for those who had enrolled in university. Finally, *Student Loan Interest (to date)* ≥ 0 is the same variable, conditional on claiming a positive amount of interest paid.

Late period decisions are analyzed using data from individuals aged $23 - 28$ over the years $2000 - 2011$. These variables are summarized in Table 5. Three primary outcomes

[28] For more information see `http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=4107`.

[29] Individuals in Canada are required to file both federal and provincial taxes and are thus required to file with a province of residence in each year.

[30] See the appendix for details.

are analyzed: the university graduation decisions, whether to move out of provinces, and whether to move out of Atlantic Canada. *University Graduate* is defined in the same way as before. It is a binary outcome defined for those who had ever enrolled in university. It is set equal to one for those who had been enrolled in university for at least 24 months, and set to zero otherwise. *Moved Province* is a binary indicator for whether an individual had moved out of province, which is analyzed only for university enrollees. The variable is set to one in the year an individual moved out of province, and any subsequent year. The variable is set to zero otherwise. *Moved From Atlantic* is similarly defined. It is set to one in the year in which an individual moves from the Atlantic Region, and any subsequent year. It is set to zero otherwise.

A secondary dataset used for analysis is the confidential version of the Survey of Labour and Income Dynamics (SLID), which is conducted by Statistics Canada. In contrast to the LAD, the SLID directly asks respondents about education outcomes. Thus, if consistent estimates are found in both samples, we can have faith in the imputation process used in the LAD analysis. The SLID surveys roughly 60,000 individuals each year, with individuals surveyed for six consecutive years.[31] Data is recorded at the individual level. The sample used is observations from the years $2000-2010$. The sample for the first part of the analysis is individuals aged $18-23$ from all provinces, except for Quebec. The sample for the second part of the analysis is individuals aged $23-28$,from the Atlantic provinces.

The analyzed early period variables from the SLID are summarized in Table 7. In general, these variables are constructed to be similar to the variables analyzed in the LAD. *University Ever* is a binary indicator for whether an individual has ever enrolled in university. *University Dropout* is a binary indicator for whether an individual has ever dropped out of university. This variable is set equal to one if the individual had ever enrolled in university, is not currently enrolled in university, and is not a university graduate. *College Ever* and *College Dropout* are similarly defined for individuals in college. Finally, *Post Sec. Ed. Ever* and *Post Sec. Ed. Dropout* are defined in the same way if the individual had enrolled(dropped out of) college or university.

Table 9 summarizes the late period variables from the SLID. These variables analyze decisions that would have been made by individuals at a slightly later stage of life. *Same Prov. as High School* is a binary indicator of whether the individual currently lives in the same province as where they attended high school. This variable is set equal to one for individuals in the same province, and set equal to zero otherwise. The variable is also summarized for individuals who had attended university, shown in the second row of the table.

[31] Unfortunately, with the SLID being a survey, there is some attrition. Survey weights are used to correct for attrition.

University Graduate is a binary indicator for whether an individual had graduated from university. Similarly, *College Graduate* is a binary indicator for whether an individual had graduated from college, and *Post Sec. Ed. Graduate* an indicator for college or university graduation.

6 Results

LAD – Early Period Decisions

Results from the early period LAD regressions can be found in Table 4.[32] The results from the regression suggest that the programs had some impacts. In general, the likelihood of individuals enrolling in university was unchanged. The variable *Ever Enrolled in University* is both very small in magnitude and highly insignificant. Similarly, the coefficient on *Enrolled in University for 2 years* is not significant. *University Dropout* is estimated to decrease by a relatively large 4.1%, and the coefficient has a cluster robust p-value of 0.021. It is not too surprising that the coefficient on *University Graduate* is not significant, as the sample here is individuals 18–23. Finally, both of the *Student Loans Interest* variables suggest that student loan interest has decreased. These coefficients are both economically and statistically significant. The sample for these variables is individuals who had enrolled in university, aged 18–23, many of whom will still be in school.

LAD – Late Period Decisions

Table 6 shows the estimates of the late period decisions using the LAD. The results show that the likelihood of graduating increased for both the combined sample, and for males. The coefficient for *University Graduate* is estimated to increase by 0.034% for the full sample, and 0.062% for males, both of which have cluster robust p-values below 5%.[33] This suggests that the programs were able to increase educational attainment. The estimates on migration decisions suggest that migration decisions were not affected. In the combined sample, the estimates for both Moved Province and *Moved From Province* are highly statistically significant. The estimate for *Moved Province* for males has a cluster robust p-value of 0.02, with a coefficient of 0.029. This suggests that university educated males were slightly more likely to leave the province when the programs were in operation. The other coefficients for the migration variables are statistically insignificant. Taken together, these estimates suggest that the programs were ineffective in retaining university educated individuals within province. This is consistent with the evidence given by the Nova Scotia government when

[32]The bootstrap p-values are forthcoming.
[33]The wild cluster p-values are forthcoming.

it cancelled its retention program.

SLID – Early Period Decisions

The results of the early period analysis using the SLID data can be found in Table 8. The results for the combined sample are displayed in the top of the table, results for males in the middle, and results for females in the bottom. In general, the results for males and females and the combined sample are quite similar, with the majority of estimates being quite statistically insignificant. One notable exception is the coefficient for *University Dropout*, which is estimated to decrease by -0.049 in the combined sample. This estimate is highly statistically significant using cluster robust standard errors, and significant at the 10% level using the 6-point bootstrap p-value. Recall that the under-rejection of the bootstrap suggests that bootstrap p-values are over stated. Interestingly, when the impacts are estimated separately by gender, the estimate for males is both larger and of greater statistical significance than for females. In fact, using the bootstrap p-values the coefficients on *University Dropout* for the combined sample and for males are the only coefficients significant at the 10% level. This suggests that the programs did not have any impact on university or college enrollment, as none of these estimates is remotely statistically significant. The procedure used for analysis is quite important as using cluster robust standard errors would lead to the conclusion that many of these coefficients are significant at the 1% level, while the bootstrap p-values would suggest otherwise.

SLID – Late Period Decisions

The results of the late period analysis using the SLID can be found in Table 10. These estimates are meant to measure the extent to which the programs influenced graduation rates and migration decisions. The most notable result from the table is that the estimated coefficients on the *Same Province* variables are negative in all three panels. This is especially interesting as the stated intention of the programs was to keep recent graduates in province. The majority of the coefficients are not remotely statistically significant when using the enumerated 6-point p-values for inference. The one exception is the coefficient on *Same Province* for males, who had attended university, which is significant at the 5% level. This suggests that these individuals were less likely to stay in province when the graduate retention programs were on offer. The lack of statistical significance on the other coefficients suggests that these individuals were not any more likely to graduate from either college or university. Again, using cluster robust standard errors would lead to incorrect inferences, when compared to the enumerated 6-point p-values.

6.1 Robustness

To check the robustness of these results, two sets of analysis are performed. The first analyzes a set of data collected from universities, at the aggregate level, on student retention. The second is a series of 'falsification tests' using the LAD data.[34]

Macleans

As a robustness check I have collected the historical retention rate from the Maclean's Magazine University Rankings.[35] The university ranking data contains self reported statistics from nearly 50 universities in Canada. The annual nature of the rankings allows for the construction of a panel, for this paper I have a panel covering the years 2002 through 2013.

The Maclean's variable measures the retention rate for each university in the sample. The retention rate reports the proportion of first year full time students who re-enroll at the same university in the following academic year. This variable is a good proxy for overall graduation rates as the majority of individuals who do not finish university leave before their second year of university.[36] The estimates in Table 11 are presented as both pooled estimates and estimates on a province by province basis.

The estimated impact of the graduate retention programs shows that the retention rate has increased by a reasonably large 2.315 percentage points. However, the statistical significance of this estimate is somewhat wanting, as the 6-point bootstrap p-value is 14.3%. The province-specific coefficients in the bottom of the table are interesting, as they suggest that the programs in Manitoba and Saskatchewan had much larger impacts than the programs in Nova Scotia and New Brunswick. Neither the cluster-robust p-values nor the 6-point p-values are reliable for these coefficients as only one group is 'treated'.[37]

7 Conclusions

Several provinces in Canada have introduced generous tax credits to increase the local stock of college educated individuals. Although the programs were designed to keep individuals in a province after graduation they explicitly encourage graduation. The impact of these programs on a variety of education and migration decisions is evaluated using both administrative and survey data. The programs have no discernible impact on college enrollment. College dropout rates decrease by 4.1% for individuals aged 18-23. College graduation rates increase by 3.4% for individuals aged 23-28 from the Atlantic provinces. Finally, the

[34] The results from this falsification exercise are forthcoming.

[35] I'm very appreciative of Mary Dwyer from Maclean's for providing me with past editions of the Maclean's rankings.

[36] See the Globe and Mail focus on 'Our Time to Lead: Education' from October 6, 2012.

[37] See MacKinnon and Webb (2014) for more details.

programs had no impact on out of province migration for recent graduates.

From a policy perspective, the implications of this analysis are somewhat ambiguous. The credits were introduced to discourage interprovincial migration. While it appears the credits did not change the pattern of migration, they did decrease university dropout rates. If those individuals who did not dropout go on to graduate, then the programs may have met the goal of increasing the average education level in the province. It is possible that the programs are superior to the Merit Scholarships in that they both offer financial relief for local students, and directly encourage individuals to graduate. Past research has also found the Merit programs to be distortionary. Cohodes and Goodman (2014) find college completion rates to decrease. Sjoquist and Winters (2013a) find that merit scholarships decrease the likelihood that a student majored in a STEM field. Future research should examine whether the retention credits had similar distortionary effects.

The credits are also perhaps a better alternative to in-state tuition, as they effectively reduce the net tuition for those who stay in province. Aghion et al. (2009) argue that investments in educational institutions in states far away from the technology frontier tend to benefit states closer to the frontier, as graduates of said institutions tend to migrate toward the frontier. Thus, it is perhaps preferable to offer reduced tuition for those who stay in province, as opposed to those who attend school in province. Whether this change would be regressive in its incidence is beyond the scope of this paper, though it is encouraging that the credits reduced the amount paid in student loan interest. While the programs appear to encourage those enrolled in university to finish their education, the costly nature of these programs makes further experimentation difficult.

References

Acemoglu, Daron, and Joshua Angrist (2001) 'How large are human-capital externalities? evidence from compulsory-schooling laws.' In 'NBER Macroeconomics Annual 2000, Volume 15' NBER Chapters (National Bureau of Economic Research, Inc) pp. 9–74

Aghion, P., L Boustan, C Hoxby, and J Vandenbussche (2009) 'The causal impact of education on economic growth: Evidence from u.s.'

Arendt, Jacob Nielsen (2013) 'The effect of public financial aid on dropout from and completion of university education: evidence from a student grant reform.' *Empirical Economics* 44(3), 1545–1562

Attewell, Paul, Scott Heil, and Liza Reisel (2011) 'Competing explanations of undergraduate noncompletion.' *American Educational Research Journal* 48(3), 536–559

Bailey, Martha J., and Susan M. Dynarski (2011) 'Gains and gaps: Changing inequality in u.s. college entry and completion.' Working Paper 17633, National Bureau of Economic Research

Belman, Dale, and John S. Heywood (1991) 'Sheepskin effects in the returns to education: An examination of women and minorities.' *Review of Economics and Statistics* 73(4), 720 – 724

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004) 'How much should we trust differences-in-differences estimates?' *Quarterly Journal of Economics* 119(1), 249–275

Bureau of Labour Statistics, U.S. Department of Labor. (2012) 'Occupational outlook handbook, 2012-13 edition, projections overview.' Online

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2008) 'Bootstrap-based improvements for inference with clustered errors.' *Review of Economics and Statistics* 90(3), 414–427

Carter, Andrew V., Kevin T. Schnepel, and Douglas G. Steigerwald (2013) 'Asymptotic behavior of a t test robust to cluster heterogeneity.' Technical Report, University of California, Santa Barbara

Cohodes, Sarah R., and Joshua S. Goodman (2014) 'Merit aid, college quality, and college completion: Massachusetts' adams scholarship as an in-kind subsidy.' *American Economic Journal: Applied Economics* 6(4), 251–85

Deming, David, and Susan Dynarski (2009) 'Into College, Out of Poverty? Policies to Increase the Postsecondary Attainment of the Poor.' NBER Working Papers 15387, National Bureau of Economic Research, Inc

Denny, Kevin J., and Colm P. Harmon (2001) 'Testing for sheepskin effects in earnings equations: evidence for five countries.' *Applied Economics Letters* 8(9), 635–637

DiNardo, John Enrico, and David S. Lee (2011) 'Program evaluation and research designs.' In *Handbook of Labor Economics,* ed. O. Ashenfelter and D. Card, 1 ed., vol. 4A (Elsevier) chapter 05, pp. 463–536

Donald, Stephen G, and Kevin Lang (2007) 'Inference with difference-in-differences and other panel data.' *Review of Economics and Statistics* 89(2), 221–233

Dowd, Alicia C (2004) 'Income and financial aid effects on persistence and degree attainment in public colleges.' *Education Policy Analysis Archives* 12(21), n21

Essaji, A., and C. Neill (2010) 'Tax expenditures vs budgetary expenditures for canadian post-secondary education.' Technical Report, Laurier Centre for Economic Research & Policy Analysis

Ferrer, Ana, and W. Craig Riddell (2008) 'Education, credentials, and immigrant earnings.' *Canadian Journal of Economics* 41(1), 186–216

Fitzpatrick, Maria D., and Damon Jones (2012) 'Higher education, merit-based scholarships and post-baccalaureate migration.' Working Paper 18530, National Bureau of Economic Research, November

Garibaldi, Pietro, Francesco Giavazzi, Andrea Ichino, and Enrico Rettore (2012) 'College cost and time to complete a degree: Evidence from tuition discontinuities.' *Review of Economics and Statistics* 94(3), 699–711

Groen, Jeffrey A. (2011) 'Building knowledge stocks locally: Consequences of geographic mobility for the effectiveness of state higher education policies.' *Economic Development Quarterly* 25(4), 316–329

Gunnes, Trude, Lars J. Kirkebøen, and Marte Rønning (2013) 'Financial incentives and study duration in higher education.' *Labour Economics* 25(0), 1 – 11. European Association of Labour Economists 24th Annual Conference, Bonn, Germany, 20-22 September 2012

Hungerford, Thomas, and Gary Solon (1987) 'Sheepskin effects in the returns to education.' *Review of Economics and Statistics* 69(1), 175 – 177

Jepsen, Christopher, Kenneth Troske, and Paul Coomes (2014) 'The labor-market returns to community college degrees, diplomas, and certificates.' *Journal of Labor Economics* 32(1), 95–121

Johnson, David R, and Fiona T Rahman (2005) 'The Role of Economic Factors, Including the Level of Tuition, in Individual University Participation Decisions in Canada.' *Canadian Journal of Higher Education* 35(3), 101–127

MacKinnon, James G., and Matthew D. Webb (2014) 'Wild bootstrap inference for wildly different cluster sizes.' Mimeo Queen's University

Magazine, Maclean's (2014) 'The rankings: this year's best.'

Malamud, Ofer, and Abigail Wozniak (2012) 'The Impact of College on Migration: Evidence from the Vietnam Generation.' *Journal of Human Resources* 47(4), 913–950

McGuigan, Marty, Sandra McNally, and Gill Wyness (2012) 'Student awareness of costs and benefits of educational decisions: Effects of an information campaign.' CEE Discussion Papers 0139, Centre for the Economics of Education, LSE, August

Moretti, Enrico (2004) 'Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data.' *Journal of Econometrics* 121(1-2), 175–212

Morissette, René, Ping Ching Winnie Chan, and Yuqian Lu (forthcoming) 'Wages, Youth Employment, and School Enrollment: Recent Evidence from Increases in World Oil Prices.' *Journal of Human Resources*

Mullan, F (1999) 'The muscular samaritan: the national health service corps in the new century.' *Health Affairs* 18(2), 168–175

National Center for Education Statistics (2012) 'Digest of education statistics - table 376.' Online

Neill, Christine (2013) 'What you don't know can't help you: Lessons of behavioural economics for tax-based student aid.' *CD Howe Institute Commentary*

Oreopoulos, Philip, and Uros Petronijevic (2013) 'Making college worth it: A review of research on the returns to higher education.' Working Paper 19053, National Bureau of Economic Research

Reamy, Jack (1994) 'Physician recruitment in rural canada: Programs in new brunswick, newfoundland, and nova scotia.' *The Journal of Rural Health* 10(2), 131–136

Riddell, W. Craig, and Xueda Song (2011) 'The impact of education on unemployment incidence and re-employment success: Evidence from the U.S. labour market .' *Labour Economics* 18(4), 453 – 463

Sartarelli, Marcello (2011) 'Compulsory and college education: A non-exhaustive survey of schooling choices and related outcomes.' Paper presented at the 2012 Annual Meeting of the Canadian Economics Association

Schneider, Mark, and Lu Michelle Yin (2012) 'Completion matters: The high cost of low community college graduation rates.' *AEI Education Outlook*

Shapiro, Jesse M. (2006) 'Smart cities: Quality of life, productivity, and the growth effects of human capital.' *The Review of Economics and Statistics* 88(2), pp. 324–335

Sjoquist, David L., and John V. Winters (2013a) 'State Merit-Aid Programs and College Major: A Focus on STEM.' IZA Discussion Papers 7381, Institute for the Study of Labor (IZA), May

— (2013b) 'The effects of HOPE on post-college retention in the Georgia workforce.' *Regional Science and Urban Economics* 43(3), 479–490

Stinebrickner, Todd, and Ralph Stinebrickner (2013) 'Academic performance and college dropout: Using longitudinal expectations data to estimate a learning model.' Working Paper 18945, National Bureau of Economic Research

The Saskatchewan Labour Market Commission's (2009) 'Right people, right place, right time - saskatchewan's labour market strategy.' Technical Report, The Saskatchewan Labour Market Commission

Turner, Sarah (2004) 'Going to College and Finishing College.Explaining Different Educational Outcomes.' In 'College Choices: The Economics of Where to Go, When to Go, and How to Pay For It' NBER Chapters (National Bureau of Economic Research, Inc) pp. 13–62

Webb, Matthew D. (2014) 'Reworking wild bootstrap based inference for clustered errors.' Working Papers 1315, Queen's University, Department of Economics

Winters, John V. (2014) 'The production and stock of college graduates for u.s. states.' Technical Report, Oklahoma State University

Table 1: Summary of the various Graduate Retention Programs

	SK	**MB**	**NS**	**NB**
maximum amount	20k	25k	15k	20k
rebate per year	10%,20%	4k, 10%	2.5k	4k
NPV ($000) @ 5%	16.9	14.1	13.3	12.6
refundable credit	Y*	N	N	N
roll-over credit	N*	Y	N	Y
eligibility duration	7	10	6	20
application req.	N	N	Y	Y
tuition based	Y	Y	N	Y
tuit. % refunded	100%	60%	-	50%
program costs	35m	34m	25m	

Notes: * reflects the change in 2012 that Saskatchewan announced, where the credit was no longer refundable, but would instead roll-over from one year to the next. The NPV calculation assumes a 5% discount rate, sufficient earnings to get the maximum credit in all years, and $22,663 in tuition paid in earning a 4 year B.A. degree from Queen's University.

Table 2: Monte Carlo Results

	CRVE	2-point boot	2-point enum	2-point enum	6-point boot	6-point enum	6-point enum
5% rej. freq.	21.552%	1.612%	9.108%	0.000%	1.828%	1.844%	1.752%
5^{th} pctile c.v.	0.005	0.120	0.000	0.125	0.115	0.116	0.117

Notes: Results from 25,000 Monte Carlo simulations.

Table 3: LAD - Summary statistics - Early Period

	GRP		Not GRP	
	Pre	Post	Pre	Post
Enrolled in University	.149	.229	.108	.146
Ever Enrolled in University	.448	.372	.39	.278
Enrolled in University for 2 years	.751	.607	.788	.583
University Graduate	.622	.417	.639	.348
University Dropout	.309	.27	.29	.316
Student Loan Interest (to date)	220	20	110	10
Student Loan Interest (to date) > 0	300	500	920	270

Table 4: LAD - Regression Results - Early Period

	Coeff.	p-value	Obs
Ever Enrolled In University	-0.00098	0.975	458692x
Enrolled in University for 2 years	0.015444	0.452	171163x
University Graduate	0.017156	0.156	171163x
University Dropout	-0.04111	0.021	171163x
Student Loan Interest (to date)	-15.7603	0.001	171163x
Student Loan Interest (to date) > 0	-42.8224	0.029	10208x

Table 5: LAD - Summary Statistics - Late Period

	GRP		Not GRP	
	Pre	Post	Pre	Post
Female	0.512	0.5	0.496	0.5
University Graduate	0.587	0.652	0.52	0.538
Moved	0.458	0.322	0.529	0.32
Moved From Atlantic	0.423	0.301	0.497	0.302
Obs	197000	86800	101400	22000

Table 6: LAD - Regression Results - Late Period

Males and Females	Coeff	s.e.	p-value
University Graduate	0.034	0.010	0.039
Moved Province	0.009	0.018	0.662
Moved From Atlantic	0.001	0.004	0.917
Males			
University Graduate	0.062	0.016	0.030
Moved Province	0.029	0.007	0.020
Moved From Atlantic	0.013	0.004	0.062
Females			
University Graduate	0.014	0.014	0.380
Moved Province	-0.011	0.026	0.707
Moved From Atlantic	-0.008	0.004	0.144

Note: Treatment Defined by Year when 23

Table 7: SLID - Summary statistics - All Provinces

	GRP		Not GRP	
	Pre	Post	Pre	Post
University Ever	0.442	0.404	0.397	0.384
University Dropout	0.089	0.022	0.090	0.037
College Ever	0.688	0.634	0.649	0.635
College Dropout	0.111	0.061	0.119	0.051
Post Sec. Ed. Ever	0.319	0.280	0.349	0.357
Post Sec. Ed. Dropout	0.129	0.113	0.135	0.072

Table 8: SLID - School Enrollment and Dropout Decisions

	Males and Females				
	Coeff.	**OLS**	**CRVE**	**6-point Boot**	**Obs**
University Ever	-0.007	0.741	0.840	0.884	37848
University Dropout	-0.049	0.004	0.002	0.087	13226
College Ever	0.006	0.760	0.781	0.890	37964
College Dropout	-0.003	0.822	0.517	0.660	24396
Post Sec. Ed. Ever	0.016	0.419	0.422	0.613	37588
Post Sec. Ed. Dropout	0.046	0.054	0.002	0.111	13728
	Males				
	Coeff.	**OLS**	**CRVE**	**6-point Boot**	**Obs**
University Ever	0.036	0.226	0.255	0.408	18592
University Dropout	-0.075	0.007	0.004	0.092	5269
College Ever	0.032	0.269	0.139	0.351	18654
College Dropout	-0.015	0.520	0.538	0.712	10877
Post Sec. Ed. Ever	0.000	0.994	0.994	0.993	18452
Post Sec. Ed. Dropout	0.030	0.389	0.352	0.548	6671
	Females				
	Coeff.	**OLS**	**CRVE**	**6-point Boot**	**Obs**
University Ever	-0.053	0.073	0.284	0.395	19256
University Dropout	-0.031	0.145	0.316	0.431	7957
College Ever	-0.028	0.277	0.385	0.545	19310
College Dropout	0.003	0.884	0.882	0.911	13519
Post Sec. Ed. Ever	0.023	0.394	0.436	0.621	19136
Post Sec. Ed. Dropout	0.049	0.111	0.214	0.332	7057

Notes: Treatment Defined by Year when 18

Table 9: SLID - Summary statistics - Atlantic Provinces

	GRP		Not GRP	
	Pre	Post	Pre	Post
Same Prov as High School	0.776	0.766	0.690	0.719
" " |Attended University	0.765	0.757	0.676	0.740
University Graduate	0.354	0.275	0.317	0.243
College Graduate	0.583	0.542	0.585	0.553
Post Sec. Ed. Graduate	0.543	0.474	0.547	0.468

Table 10: SLID - Migration and Graduation Decisions

	Males and Females				
	Coeff.	**OLS**	**CRVE**	**6-point Enum**	**Obs**
Same Prov as High School	-0.062	0.190	0.343	0.371	9324
“ " |Attended University	-0.113	0.031	0.053	0.172	7324
University Graduate	0.049	0.382	0.457	0.757	3959
College Graduate	0.059	0.307	0.515	0.652	4958
Post Sec. Ed. Graduate	0.061	0.168	0.314	0.558	7324
	Males				
	Coeff.	**OLS**	**CRVE**	**6-point Enum**	**Obs**
Same Prov as High School	-0.054	0.404	0.388	0.359	4467
“ " | Attended University	-0.163	0.023	0.027	0.047	3273
University Graduate	0.117	0.156	0.333	0.350	1599
College Graduate	0.088	0.269	0.148	0.211	2401
Post Sec. Ed. Graduate	0.114	0.085	0.210	0.166	3273
	Females				
	Coeff.	**OLS**	**CRVE**	**6-point Enum**	**Obs**
Same Prov as High School	-0.084	0.178	0.428	0.434	4857
“ " | Attended University	-0.090	0.187	0.284	0.234	4051
University Graduate	0.012	0.869	0.572	0.839	2360
College Graduate	0.025	0.748	0.853	0.806	2557
Post Sec. Ed. Graduate	0.002	0.968	0.972	0.987	4051

Notes: Treatment Defined by Year when 23

Table 11: Maclean's - Percentage of Students Returning for 2nd Year

	Coeff.	**OLS**	**CRVE**	**6-pt Boot**
All	2.315	0.021	0.053	0.143
SK	4.594	0.000	0.000	0.622
MB	4.006	0.000	0.000	0.530
NS	2.141	0.000	0.001	0.320
NB	-1.009	0.043	0.074	0.550

Notes: Estimates Clustered by Province

8 Appendix

The procedure used by Morissette, Chan and Lu (forthcoming) (MCL, henceforth) to estimate university enrollment is described in detail in the appendix of their paper. This procedure borrows heavily from theirs.[38] Since 1999 the tax code has the following three variables which allow for inference of university enrollment rates:

- tuition fees for self
- educational deduction for full-time students
- educational deduction for part-time students

MCL use these variables to construct enrollment rates, the procedure is as follows:

First, I take the total education deduction for full time students and divide by the maximum amount allowed per month by the Canada Revenue Agency.[39] This gives the total number of months the individual was enrolled full-time in college or university that year.

Second, a similar calculation is performed for the part-time months.

Third, an estimate of the months of full-time-equivalent enrollment is derived. Full-time months are counted as 1.0 months and part-time months are counted as 0.6 months.

Fourth, the amount claimed in tuition is divided by the number of months estimated in step 3. This yields an estimated tuition per month.[40]

Fifth, the monthly average tuition amount is multiplied by 8, to compute a full-time-equivalent annual amount.

Sixth, the annual amount from step 5 is compared to $0.8-2.0$ times the annual provincial average undergraduate tuition amount.[41] If the annual amount is within the interval, then the variable "university enrolled" is set equal to 1, it is set equal to 0 otherwise.[42]

[38] I am grateful to René Morissette for providing me with SAS code to implement the procedure.

[39] The maximum amounts come from either Morissette, Chan and Lu (forthcoming) or Neill (2013).

[40] To correctly estimate the average tuition, the amount reported is converted into a full-time-equivalent tuition amount.

[41] The values used are two year averages, owing to the overlap of the school year and the tax year. The tuition values used from 1977 – 2011 come from the following sources: the values from 1977 – 2000 come from Johnson and Rahman (2005), I thank David R. Johnson for providing me with these values; 2001 – 2008 come from MCL; 2009 – 2011 from Statistics Canada Table B.2.9 http://www.statcan.gc.ca/pub/81-582-x/2014001/tbl/tblb2.9-eng.htm.

[42] Unfortunately, for the purposes of this study, the tax system does not differentiate between college and university enrollment. The MCL procedure uses the fact that university tuition is much more costly than college tuition to infer university attendance. In 2013 average undergraduate university tuition in Canada was $5,772 (see statscan table referenced above), while average college tuition in Ontario was $1,900 (http://www.tcu.gov.on.ca/pepg/audiences/colleges/costs_coll.html#tuition).

Building on the MCL procedure, I take advantage of the panel nature of the dataset to construct additional variables. I sum the months variable from step three over all the years in the dataset, to estimate the total months spent in university. I then compare this variable to a number of key thresholds. For "persistence", I estimate whether the individual completed two or more years of university, by calculating whether the total number of months exceeded 16.

Similarly, I estimate whether someone graduated from university by comparing the total months variable to 24. University programs are typically 32 months in length in most of Canada, but programs in Quebec and non-honours undergraduate programs can be 24 months in length.[43]

Finally, university dropout is set equal to 1 if an individual was a) not already a university graduate, b) had previously been enrolled in university, and c) was not currently enrolled in university.

[43] The majority of dropouts occur early in university careers. Stinebrickner and Stinebrickner (2013) report that of those who dropped out –of one particular college – 40% dropped out in year 1, 34.4% in year 2, and 25.1% in year 3. Similarly, the Maclean's graduation rate is on average 86.5% of the retention rate, suggesting that the largest fraction of dropouts leave after the first year. Thus the majority of those who finished 24 months of university will be university graduates.