UPPSALA UNIVERSITET

Department of Economics
Working Paper 2019:8

Measuring the Effect of Student Loans on the College Dropout Rate

Alex Solís

Department of Economics
Uppsala University
Box 513
751 20 Uppsala
Sweden

Working Paper 2019:8
August 2019
ISSN 1653-6975

Measuring the Effect of Student Loans on the College Dropout Rate

Alex Solís

Papers in the Working Paper Series are published on internet in PDF formats.

Measuring the Effect of Student Loans on the College Dropout Rate

Alex Solís*

August 19, 2019

Abstract

Most governments around the world offer student loans to help disadvantaged students to enroll in college to reduce the attainment gap between rich and poor. However, we know little about the consequences of these loans. The reduction of the gap depends not only on initial enrollment but also on the dropout rate before graduation. This paper shows how the availability of loans affects the dropout rate in college. Two programs in Chile assign loans based on a cutoff in the national college admission test, enabling a regression discontinuity design. The analysis uses on students who were not induced by the loan to enroll in the first year. I show that access to loans reduces the dropout rate by 25 percentage points and is highly persistent over time (up to the fifth year after initial enrollment). At the cutoff, access to loans allows eliminating the differences in the dropout rate by family income. Finally, I find that students are not sensitive to tuition costs when loans are available.

*I am grateful to David Card for his contributions to this paper. Special thanks to Susan Dynarski and Eva Mörk for their comments; Also Marcelo Lopez, Rodrigo Rolando and Juan Salamanca from SIES at the Ministry of Education of Chile; Gonzalo Sanhueza, Daniel Casanova, and Humberto Vergara from the Catholic University of Concepción; and Jorge Campos and Felipe Gutierrez from the INGRESA commission for providing the data. The paper also benefited from comments from seminar participants at Helsinki University, Lund University, and the Nordic Labor Summer Institute in Bergen. Solís: Department of Economics, Uppsala University alex.solis@nek.uu.se

1 Introduction

Many countries around the globe spend a significant amount of resources on financial aid for college in an attempt to close the educational attainment gap between individuals from rich and poor backgrounds. Despite their importance, we do not have a complete understanding of the consequences of different types of aid (grants, scholarships, tuition waivers, subsidized loans, and unsubsidized loans) for the different margins involved. Most financial aid affects initial enrollment, but these policies may not reduce the attainment gap if the students who are induced to enroll, fail to graduate. The effectiveness of aid policies depends on the effects on both initial enrollment and persistence until graduation. This paper tries to increase our understanding of the effects of students loans on the college dropout rate.

Estimating the effect of loans on the dropout rate is challenging because of at least three factors. First, loan access is usually correlated with unobserved variables, such as family income, wealth, and background, which affect dropout directly and create an omitted variable bias in observational studies. Second, the study of persistence cannot use an exogenous variation on aid that affects initial enrollment, because that variation affects sample participation (students need to be enrolled to persist or drop out) producing biased estimates. Therefore, the researcher needs an additional source of variation in aid access that does not affect enrollment. Third, the analysis of the dropout rate requires a significant amount of data: it requires data from multiple institutions to avoid misclassification of students who switch institutions, as well as information from several periods since the dropout rate is sensitive to measures over time.

In this paper, I address these three problems using a natural experiment in Chile that offers college loans to students. First, to be eligible, students from the four poorest income quintiles need to score above a cutoff (475 points) on the national college admission test, enabling a regression discontinuity design and providing an unbiased estimation of the causal effect of college loans on persistence.[1] Second, to avoid sample selection, I follow a small sample of students who enrolled

[1]In a small vicinity around the eligibility cutoff, students are as good as randomly assigned to loan access (as in Lee

without being eligible for loans in the first year but participated in the assignment of loans in the second year of college. Third, I construct a unique panel of ten years of administrative records to track the population of students in all higher education institutions in the country to avoid misclassification of students who change educational institutions.

I find that second-year ineligible students are twice as likely to drop out than those eligible for loans. The dropout rate drops from 47% to 23% at the cutoff. The results are robust to different specifications and bandwidths in the estimation. The 24 percentage point difference persists over the third, fourth, and fifth years of college.

I propose a decomposition over time and over chosen outside option to understand the dynamics of the dropout rate. I find that most of the accumulated effect is explained by the behavior in earlier periods, suggesting that dropping out may be difficult to revert. Moreover, I show that students without access to loans move to lower-quality education in the vocational sector, where they do have access to financing.

Two alternative mechanisms can explain the effect. First, aid in the form of grants or subsidized loans implies a reduction in the initial investment cost, leading to an increase in the internal rate of return to education, which, in turn, motivates students to stay in college until graduation. Second, financial aid implies the alleviation of financial constraints that prevent persistence. To test these mechanisms, I use a second natural experiment occurring at a different cutoff. Students from the two poorest income quintiles who score more than 550 points in the PSU test are eligible for the scholarship "Beca Bicentenario" (BC hereafter), enabling a similar regression discontinuity analysis.[2] Given that the BC scholarship funds the same amount as the loan to cover tuition,[3] students who cross the cutoff evidence a reduction in the cost they need to fund for their education. The estimated effect at this new cutoff is one percentage point change and is not significantly different from zero, suggesting that students are not sensitive to tuition levels when they have credit access

(2008)). Thus, the difference in persistence at the cutoff can be attributed to credit restrictions.

[2] Similar analyses have been done by Solis (2013) and Solis (2017)

[3] The maximum combined benefit is determined by the government and is referred to as the *Reference Tuition*. On average, it covers 90% of tuition costs.

in this context.

Finally, I examine the educational attainment gap by family income. I find that the effects are strongest and most stable over time for the lowest two income quintiles, while for the third and fourth quintiles, the effect vanishes after the fourth year. The results indicate that access to loans significantly reduces the persistence gap between the bottom and the top income quintiles. The gap in the dropout rate among barely ineligible students is 20 percentage points after five years of initial enrollment. Among those who are barely eligible, the gap disappears in the same period. Additionally, I explore the effects by gender; and find a stronger effect for women. This evidence may suggest a higher barrier for women to finance their education.

This paper contributes to the literature that studies the relationship between student aid and persistence, in which loans are considered one of few policy tools. Most of the literature study grants using US data. Some papers address the challenges in the estimation in different contexts, finding mostly positive and significant effects of grants on persistence (for example, DesJardins et al. (2002), Dynarski (2003), Bettinger (2004), Singell (2004), Castleman and Long (2016), Denning (2018) and Denning et al. (Forthcoming)).[4,5] Also studying grants in the U.S., other studies find some negative effects on graduation, with grants diverting students towards colleges that offer more aid but have lower graduation rates (Cohodes and Goodman (2014) and Angrist et al. (2016)). Finally, using data from the U.K., Murphy and Wyness (2016) find that non-salient additional grants increase the probability of a "good-degree."[6] Most of the literature studies discrete increases in grant

[4]See Chen (2008), Goldrick-Rab et al. (2009) and Hossler et al. (2009) for a discussion of the literature on college aid.

[5]The literature uses different variables of interest that are closely related to the definition of the dropout rate used in this paper. In many papers, the variable of interest is called the persistence or retention rate, which is defined as an indicator of whether a student appears enrolled in a given year after the first year. In this paper, the dropout rate is exactly the complement of the previous definition of persistence, and we use the terms interchangeably when it is possible. In other papers, the variable of interest has been the stopout rate, which refers to an indicator of whether a student has not been enrolled in any given year.

[6]Other papers study the effects of marginal financial incentives that are conditional on student performance on post-enrollment outcomes. Randomized aid schemes provide mixed evidence. Brock and Richburg-Hayes (2006) find positive effects, while Angrist et al. (2009) find positive effects only for women. Similarly, Garibaldi et al. (2011) study how an increase in tuition after the nominal graduation time leads to a higher probability of on-time graduation for students at Bocconi University in Italy. Similar substitutions patterns are found between public schools and for-profit colleges when aid becomes restricted on for-profit colleges (Cellini (2009) and Cellini et al. (Forthcoming)).

amounts and focuses on information from one or a few institutions. Also, most of the earlier papers focus on short-run measures of dropout, while some studies use variation in initial enrollment to study persistence.

This paper contributes showing the first evidence relative to loans; making sure that the sample of students has not been induced to enroll by the variation on aid, and including national coverage in enrollment and complete longitudinal information on enrollment and aid.[7]

The second contribution of this paper is the study of the mechanisms underlying the dropping out behavior. Few of the previous studies have attempted to disentangle the mechanisms. Using survey data from Berea College in Kentucky, Stinebrickner and Stinebrickner (2008) ask students directly whether or not they wanted to borrow more while in college; they find little support for the need for more loans and no relation to persistence. Scott-Clayton (2011) studies tuition waivers that are conditional on students' performance, concluding that the main driver of the higher persistence is the increased effort and not credit constraints. This paper contributes to this small literature, showing, in this context, that changes in tuition costs do not affect persistence once loans are available, suggesting that the results are driven by credit access.

Additionally, this paper also contributes to the small literature of debt aversion– i.e., the disutility of holding debt–on college decisions. Rothstein and Rouse (2011) and Field (2009) show that holding debt induces students to change educational and labor market decisions. Relative to dropping out during college, Goldrick-Rab et al. (2012) find that additional grants increase persistence only when students receive additional grants on top of federal loans (without replacing). Nevertheless, the substitution of loans for grants, in Chile, does not affect persistence once credit access is secured, suggesting little evidence of debt aversion.

Previously, Solis (2017) tested the effect of these loan programs on whether students have ever enrolled for two years, exploiting the variation in loan access at initial enrollment. As a result, it is not possible to disentangle the enrollment and persistence effects.[8] In this paper, however, I use

[7]Evidence indicates that at this margin, the likelihood of moving to other countries to study is very low.

[8]Recently, other papers study the effects of the same policy reform on graduation but using the variation at enroll-

the second-year test score of students who were ineligible after their first attempt and retook the test. Although this additional source of variation comes at the cost of a reduced sample, the effects are robust and visible with the usual graphical analysis.

The paper proceeds as follow. Section 2 explains the context, the loan programs, and the data. Section 3 describes the estimation strategy and sample selection. Section 4 describes the main results. Section 6 revisits the education gap by family income, and Section 6.2 analyzes the effects by gender. Section 7 concludes.

2 Context, Loan Programs, and Data

Obtaining a university degree at one of Chile's 55 universities creates a significant financial burden. The share of family income that goes to pay tuition would be 50% for a family in the median of the income distribution. Moreover, most college programs are designed to last five years, and students take six, on average, to graduate. As a consequence, students rely heavily on family resources, grants, or college loans.

Eligibility rules in the financing system in Chile (loans and grants) are very simple. Most aid programs use cutoffs on the college admission test (Prueba de selección Universitaria, PSU hereafter) for given income quintiles. Students apply for all loans and aid programs at once, filling out a simple form called FUAS (Formulario Único de Acreditación Socioeconóomica). The information on the form is contrasted with official records from the tax authority, which determines the income quintile of the students. All students classified in the four lowest income quintiles are considered *pre-selected* for loans. This financial form is similar to the one suggested by Dynarski and Scott-Clayton (2013) to increase access for low-income students in the US.

The two most important financing programs are the Traditional University Loan (TUL hereafter) and the State Guaranteed Loan (SGL hereafter),[9] benefiting about 18% of all students taking the

ment (Montoya et al. (2017), Barrios Fernández (2019), and Bucarey et al. (2018)). The data used in this paper have also been used by Santelices et al. (2015) to study persistence using propensity score matching.

[9]TUL corresponds to Crédito Solidario para Universidades del CRUCH, and SGL to Crédito con Aval del Estado

college admission test each year and 50% of those who are pre-selected. Despite some differences between the loans, students from the poorest four income quintiles become eligible if they score more than 475 PSU-points.

These tuition loans finance up to a maximum amount determined by the Ministry of Education, the so-called reference tuition (Arancel de referencia), which depends on the prestige of the institution and future labor market prospects. On average, this reference amount covers about 90% of the actual tuition, and there are no loans or grants to cover other expenses, except for some subsidies for very poor students (covering lunch and transportation) or scholarships given to outstanding students. Despite TUL and SGL, students depend on family support to finance all other costs associated with obtaining a college degree. Hence, even students who are eligible for loans may be credit-constrained and unable to fund all costs. Therefore, the effects described here correspond to a lower bound of the complete elimination of such constraints.

TUL and SGL differ in two main aspects. First, TULs are granted only to students who enroll in a traditional university, while SGLs are given to students who attend any accredited university.[10] Second, the loans are managed by different organizations and have slightly different payments conditions. TULs are low-interest (2% per year) loans managed directly by universities from resources given by the government. SGLs are handled directly by commercial banks and have conditions that resemble those of conventional loans, having an interest rate close to 6%; for the period analyzed; this is similar to the interest on other loans of equivalent duration, but the interest rate was reduced to 2% in 2011 after strong protests from students.[11] According to the World Bank report on the SGL program, the loan does not have an embedded subsidy, and the contract terms should lead to a high

(CAE is the Spanish acronym).

[10]Traditional universities are institutions funded before 1980, and they receive direct government funding. All other universities are called private universities, and they receive only indirect government funding if they enroll the best students.

[11]They also differ in less critical aspects such as the grace period and the repayment scheme. Students have a grace period after graduation of two and 1.5 years for TUL and SGL, respectively. Installments in SGL are calculated as fixed amounts with a horizon of 15 to 20 years, while TUL is an income-contingent loan with the minimum repayment set at 5% of the borrower's income. More details about these two loan programs and other funding alternatives in the Chilean university system can be found in Solis (2017).

recovery (World Bank (2011) pp. 30).

Importantly, enrolled students can also become eligible for the SGL if they complete 70% of the course work.

Loans from commercial banks and other financial institutions may constitute an alternative. Nevertheless, in the conventional market, the credit requirements are higher. For example, banks offer college loans indirectly through parents, who are required to prove that they have a formal sector job and a salary above a threshold, thus excluding the vast majority of students from the poorest three income quintiles. First, the minimum salary required is above the median of the wage distribution. Second, in Chile, accounting for 30% of workers, especially those in the lower parts of the income distribution, are employed in the large informal sector, which cannot certify income.

2.1 Others grants

Other grants and scholarships are also assigned based on cutoffs on the PSU test. However, all of them use higher cutoffs and, therefore, do not confound the effect of these two loan programs.

The largest of these grants is the Bicentenario Scholarship, which covers the reference tuition (the same amount as the loans) for students from the lowest two income quintiles who score more than 550 PSU-points and enroll in traditional universities. The scholarship shares the application process through the completion of the FUAS form, and is the third most important program, just after the two types of loans described above. It benefits about 5% of all students taking the PSU each year and is given to 55% of all of the pre-selected students.

Around the scholarship cutoff, all potential beneficiaries of Bicentenario are also eligible for loans; therefore, meeting the cutoff allows students to substitute loans for the grant altogether. Specifically, at this cutoff, we can determine the importance of the implicit subsidy component of the loan, given that receiving the grant is equivalent to a reduction in the price of college.

2.2 Data

I construct a panel of administrative records for all individuals that participated in the admission system between 2007 and 2016, merging registries from three sources using national identification numbers.

The first is the registry of students who enroll to take the PSU test. It gives information on the scores (each attempt), high school GPAs, and a comprehensive set of socioeconomic characteristics, such as parents' education, school and year of graduation. The second registry corresponds to the higher education enrollment from the Ministry of Education, containing the universe of students who enroll in higher education institutions in the country. The third source is the FUAS application dataset with the income classification from the tax authority and eligibility for loans and grants from the Ministry of Education.

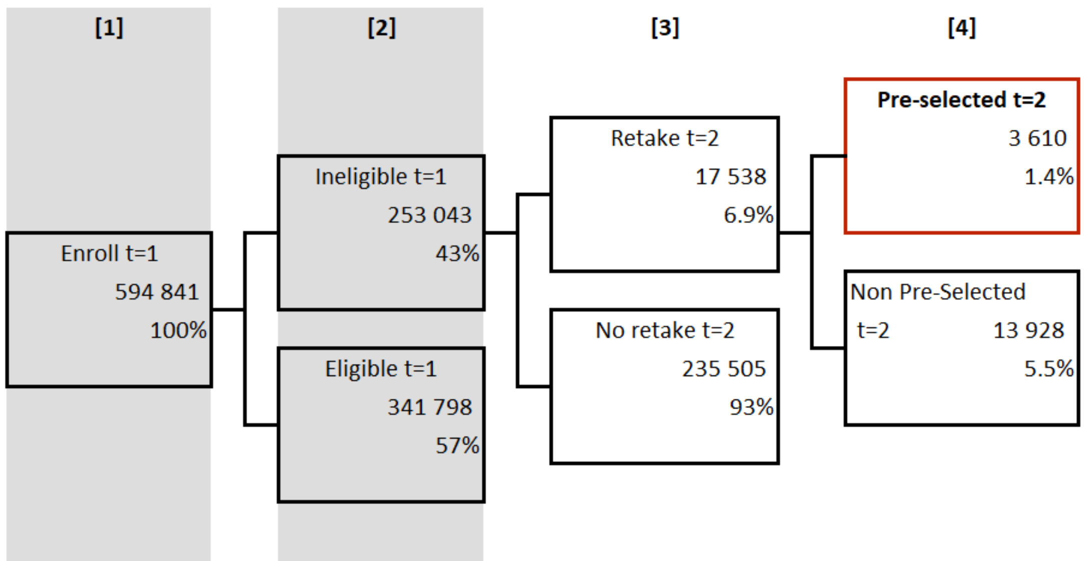
2.3 Sample

Dropping out decisions are conditional on initial enrollment; however, most of the times, aid programs that offer college loans are assigned before initial enrollment and, in most cases, the status is kept during all the years of college. As a consequence, exogenous variation in credit access assigned before enrollment cannot be used to determine the effects on dropout rates because enrollment is a decision induced by the availability of loans. The composition of the comparison group will be different. The group without access to loans (the control group) will consist only of students who enroll regardless of eligibility, whereas, the group with access to credit (the treatment group) will be formed by those who enroll regardless eligibility, and the students who enroll induced by the access to credit. Therefore, enrolled students with and without access to credit are not comparable even when the initial loan assignment is random.

I address this concern using a sample of enrolled students who are not eligible for loans at initial enrollment because they scored below the cutoff on their first PSU attempt, or did not apply for benefits in the first year. However, they have the opportunity to become eligible after their first year of college, when they take the admission test for a second time, conditional on fulfilling the eligibility requirements. Students scoring around the loan cutoff on the second attempt would be as good as randomly assigned to loans for the second year onward.

In this way, the exogenous variation in credit access affects a homogeneous sample of enrolled students and can be used to estimate the causal effects of loans on the dropout rates.

Figure 1: Sample.



Note: Column (1) shows the number of students enrolled in the first PSU attempt. Column (2) divide the population into ineligible and eligible for loans in the first attempt (t=1). Column (3) described the retaking behavior in t=2. In column (4) shows the sample of students that retaking the PSU in t=2 become pre-selected for loans in t=2.

The overall population consists of approximately 600,000 students who enroll in their first PSU attempt; thus, are subjects to drop out. The different decisions and conditions are depicted in Figure 1. I first restrict the sample to students who enrolled in college without being affected by loan availability (ineligible students in $t = 1$, upper box in column (2)). Moreover, only 6.9% of the initial population retake the test to become eligible in t=2 (upper box in column (3)). Finally, only students who complete the FUAS form and are classified in the bottom four income quintiles are affected by crossing the cutoff, reducing the sample to 3,610 students, equivalent to 1.4% of the initial population of students who enrolled. These sample restrictions highlight the difficulties faced when studying the impact of loan access on persistence. Nevertheless, despite the small sample, the unique circumstance of the Chilean case allows an arguably reliable estimation of the effects.

Table 1: Descriptive Statistics

	Population in t=1		Retakers in t=2		Sample in t=2	
	Mean	s.e.	Mean	s.e.	Mean	s.e.
Variable	(1)	(2)	(3)	(4)	(5)	(6)
PSU score in t=1	566.57	(87.27)	515.74	(86.09)	529.64	(93.91)
PSU score in t=2	600.52	(78.08)	543.91	(85.17)	543.15	(92.89)
Language score t=1	564.77	(93.06)	516.71	(91.88)	528.55	(101.39)
Math score t=1	567.57	(98.67)	514.4	(96.22)	529.66	(103.94)
High school GPA	5.83	(0.49)	5.67	(0.46)	5.71	(0.49)
Public high school	0.29	(0.45)	0.34	(0.47)	0.31	(0.46)
Voucher high school	0.49	(0.5)	0.54	(0.5)	0.52	(0.5)
Private high school	0.21	(0.41)	0.11	(0.32)	0.16	(0.36)
Age in t=0	19.37	(2.26)	18.97	(1.44)	19.38	(2.23)
Female	0.52	(0.5)	0.57	(0.49)	0.52	(0.5)
Quintile in t=1	2.64	(1.38)	2.5	(1.41)	3.34	(1.71)
Quintile in t=2	2.54	(1.37)	2.42	(1.32)	2.47	(1.16)
Quintile$_{t=2}$ = 1	0.28	(0.45)	0.33	(0.47)	0.28	(0.45)
Quintile$_{t=2}$ = 2	0.23	(0.42)	0.25	(0.44)	0.24	(0.43)
Quintile$_{t=2}$ = 3	0.18	(0.39)	0.18	(0.38)	0.21	(0.41)
Quintile$_{t=2}$ = 4	0.18	(0.38)	0.15	(0.35)	0.27	(0.44)
Quintile$_{t=2}$ = 5	0.13	(0.33)	0.09	(0.29)	-	-
Self-reported income cat.	1.17	(0.42)	1.16	(0.42)	1.21	(0.46)
Household size	4.35	(1.85)	4.37	(1.74)	4.35	(1.85)
Father years education	12.89	(3.65)	12.37	(3.58)	12.95	(3.7)
Mother years education	12.68	(3.42)	12.15	(3.37)	12.81	(3.37)
Mother house wife	0.4	(0.49)	0.43	(0.5)	0.38	(0.49)
Father in formal job	0.6	(0.49)	0.59	(0.49)	0.57	(0.49)
Mother in formal job	0.39	(0.49)	0.37	(0.48)	0.4	(0.49)
Mother College	0.22	(0.41)	0.16	(0.37)	0.22	(0.42)
Father College	0.18	(0.39)	0.13	(0.34)	0.2	(0.4)
Father dropout	0.3	(0.46)	0.32	(0.47)	0.29	(0.45)
Mother dropout	0.28	(0.45)	0.31	(0.46)	0.27	(0.44)
Observations	594,832		215,731		3,610	

Note: The first two columns describe the whole population. The next two describe students who retake the test in the following years, and the last two describe the sample used in the analysis– i.e., those who enrolled without access to loans in $t = 1$, but who were classified as pre-selected and retook the test in $t = 2$.

These students may differ from the average student in the country, in that they can enroll despite having no access to college loans in the first year. I test how comparable the students in my sample are to the overall population. However, the effects will be only internally consistent and cannot be extrapolated to general populations.

Table 1 shows descriptive statistics for three different groups: the group of students who enrolled in college after their first PSU attempt, the sample of retakers, and the sample of students used in the analysis.

The sample contains students with slightly lower ability: with lower high school GPA and test scores in both periods. However, background characteristics appear very similar among groups, especially about parental education.

3 Estimation Strategy

In a small vicinity around the cutoff, enrolled students scoring barely above are similar in all dimensions to students scoring barely below the threshold, except that they are eligible for loans for the rest of the investment period. Therefore, both groups constitute a good counterfactual to estimate the effects of credit access on dropping out decisions.[12]

I follow Imbens and Lemieux (2008) to run the following specification:

$$P(E_{2i} = 0|S; T_{2i}) = \alpha_1 + \beta_1 \cdot \mathbb{1}(T_{2i} \geqslant c) + f(T_{2i} - c) + \epsilon_{2i} \quad (1)$$

where T_{2i} corresponds to i's PSU score on the second attempt ($t = 2$). $P(E_{2i} = 0|S; T_{2i})$ is the probability of not being enrolled in $t = 2$, conditional on the PSU score T_{2i} and being part of the sample of analysis, S (described above). The treatment indicator, $\mathbb{1}(T_{2i} \geqslant c)$, takes the value of 1 when student i scores at least the eligibility cutoff, $T_{2i} \geqslant c$, in year $t = 2$,and zero otherwise. When this condition is satisfied, student i is eligible for loans for the rest of her studies. Finally, since the running variable is a good measure of ability, which is a crucial variable in the determination of the dropout rate, I control for this ability using the PSU score in $t = 2$, T_{2i}, using a flexible control

[12]The RD using the test score in $t = 2$, T_{2i}, is sharp in the second year because the policy was rigorous, and nobody who scored less than the cutoff was able to get the loans. On the other hand, scoring at least the cutoff ensured eligibility or access to loans for students in our sample.

function at each side of the cutoff, $f(T_{2i})$.[13]

In this specification, α_1 captures the average dropout rate among barely ineligible students, while β_1 captures the causal effect of having access to loans on the dropout rate for this local population.

Given that students can keep taking the test in later years, they may become eligible in future attempts. Therefore, the analysis of longer-run outcomes requires controlling for this dynamic selection into treatment. Being eligible for loans in $t > 2$ is no longer completely determined by the score on the second attempt. Nevertheless, the probability of being eligible for loans would still change discretely if some students did not succeed in scoring at least the cutoff. This situation happens when they do not retake the PSU or when they are not able to score more than the cutoff.

If this is true, being eligible for loans can be instrumented by the treatment indicator of crossing the cutoff in the second attempt, $\mathbb{1}(T_{2i} \geqslant c)$. This instrumental variable is valid because it is exogenous to the error term and is correlated with loan eligibility in other periods.

The following model captures the causal effect of ever being eligible for loans on dropping out after the second year.

$$EverEligible_{ti} = \alpha_2 + \beta_2 \cdot \mathbb{1}(T_{2i} \geqslant c) + f_2(T_{2i} - c) + \nu_{ti} \tag{2}$$

$$P(E_{ti} = 0|S; T_{2i}) = \alpha_3 + \beta_3 \cdot EverEligible_{ti} + f_3(T_{2i} - c) + \xi_{ti} \tag{3}$$

$EverEligible_{it}$ is an indicator of whether the student crosses the threshold in any $t \geqslant 2$, becoming pre-selected for loans in that period. It is an endogenous variable since students may choose to retake the test. $P(E_{ti} = 0|S; T_{2i})$ is an indicator variable that takes the value one if student i drops out in $t > 1$. As before, $f_q,\ q = 2, 3$ is a function that controls for the influence of the running variable, the PSU score on the second attempt T_{2i}.

[13]The main results use a linear control on a uniform Kernel at each side of the cutoff – i.e.,$f(S_i) = \phi_0 S_i + \phi_1 S_i \cdot \mathbb{1}(S_i \geqslant c)$. To show robustness, I present results using a third-order polynomial at each side.

3.1 Dynamic and by Destination Decomposition of the Dropout Rate

The probability of not enrolling in $t=2$, $P(E_{2i}=0|S;T_{2i})$ can be decomposed depending on the alternative chosen by the student. Whether the student decided to enroll in a vocational program, $P(V_{2i}=1|S;T_{2i})$, or whether she chose to enter the labor market, $P(M_{2i}=1|S;T_{2i})$:[14]

$$P(E_{2i}=0|S;T_{2i}) = P(V_{2i}=1|S;T_{2i}) + P(M_{2i}=1|S;T_{2i}) \tag{4}$$

Each of these probabilities on the right-hand of Equation (4) can be estimated as in Equation (1), where the dependent variable is changed accordingly.

Moreover, the probability of not enrolling in $t=3$ can be decomposed temporally to reflect the time of initial dropout: students who dropped out in $t=2$ and did not come back in $t=3$, $P(E_{2i}=0, E_{3i}=0|S;T_{2i})$; and the group that enrolled in $t=2$, but failed to register in $t=3$, $P(E_{2i}=1, E_{3i}=0|S;T_{2i})$.

$$P(E_{3i}=0|\cdot) = P(E_{2i}=1, E_{3i}=0|\cdot) + P(E_{2i}=0, E_{3i}=0|\cdot) \tag{5}$$

To simplify the notation, I will refer to these probabilities as $P(E_{2i}=j_2, E_{3i}=j_3|\cdot) = p_{3,j_2,j_3}$, where j_2 indicates the alternative chosen in $t=2$, and j_3 the decision on $t=3$. j_2 and j_3 can take the value 0, 1, v or m: whether the students enrolled in college, 0 or 1; did not enroll because they went to a vocational program, v; or did not enroll because they entered the labor market, m.

$$P(E_{3i}=0|\cdot) = p_{3,.,0} = p_{3,1,0} + p_{3,0,0}$$

Adding the decomposition by the chosen alternative I can describe the dropout in $t=3$ as follow,

[14] I assume that individuals who do not enroll in any institution go to the labor market. However, since I do not have information about employment, I cannot distinguish if the individual is working, enjoying leisure, or preparing to retake the PSU.

$$
\begin{aligned}
P(E_{3i}=0|\cdot) &= P(E_{2i}=1, E_{3i}=0|\cdot) + P(E_{2i}=0, E_{3i}=0|\cdot) \\
&= P(E_{2i}=1, V_{3i}=1|\cdot) + P(E_{2i}=1, M_{3i}=1|\cdot) \\
&+ P(E_{2i}=0, V_{3i}=1|\cdot) + P(E_{2i}=0, M_{3i}=1|\cdot)
\end{aligned} \tag{6}
$$

$$
p_{3,.,0} = p_{3,1,v} + p_{3,1,m} + p_{3,0,v} + p_{3,0,m} \tag{7}
$$

where each of the elements of the right-hand side of equation (6) or (7) can be estimated separately using (1).

Similarly, I can write $P(E_{4i}=0|\cdot) = p_{4,.,.,0}$ or $P(E_{5i}=0|\cdot) = p_{5,.,.,.,0}$, and their corresponding decompositions. For example, $p_{4,.,.,0} = p_{4,0,0,0} + p_{4,0,1,0} + p_{4,1,0,0} + p_{4,1,1,0}$, for the temporal decomposition in the fourth year (See more details about the decompositions presented in the paper in the Appendix A).

3.2 Definition of Variables

Because the history of the educational decisions that we observed is truncated after some periods (e.g., after ten years for cohort 2007), we may misclassify students who stopped for some years but planned to re-enroll later. Consequently, the estimates may vary depending on the definition of dropping out.

First, I define dropout by merely observing the enrollment in a given year. The variable $P(E_{si} = 0|S, T_{2i})$ takes the value one if somebody is not enrolled in year $s \geqslant 2$, regardless of what happens between $t = 2$ and $t = s$. This definition is agnostic about whether some students may try different strategies to survive in college, such as stopping for work and saving money.

This definition of dropout is the most used in the literature; for example, in Bettinger (2004) and Singell (2004), a dropout is defined as any student who does not matriculate in the second year after initial enrollment.

An alternative definition, $P(E_{2i} = 0 \text{ or ... or } E_{si} = 0|\cdot)$ takes the value one if a student stops out in any year between $t = 2$ and $t = s$. Therefore, it takes the value zero only if the student has been enrolled continuously for s years. Both measures of dropping out are equivalent in the second year of college, which is our main object of interest. DesJardins et al. (2002) use the time to first stop-out as the outcome variable (the first non-continuous enrollment). In their context, the main reason to use this variable is its high correlation with dropping out permanently.

In this paper, given that both definitions lead to similar conclusions, I will focus on the first.

4 Results

4.1 Validity of the RD

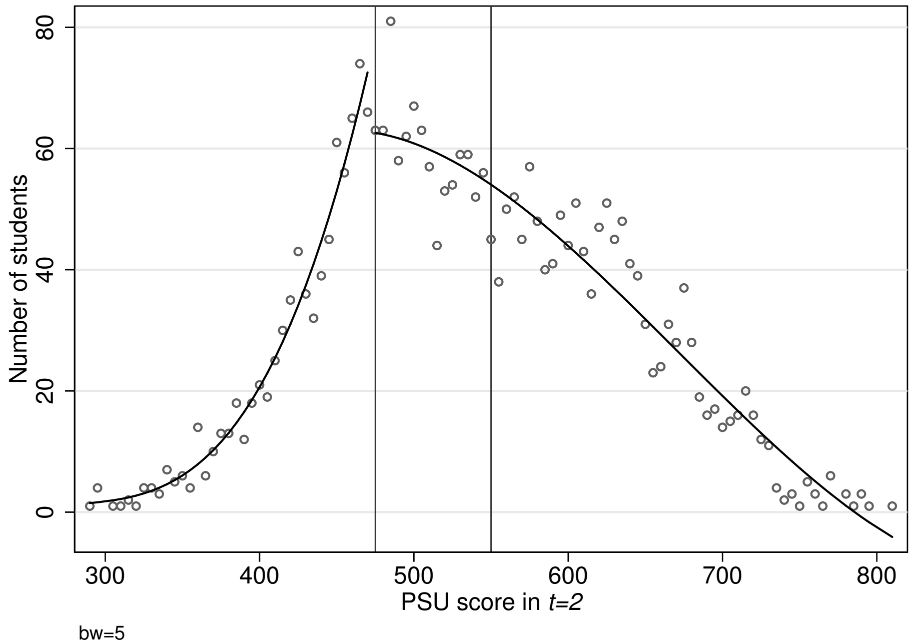
Following the standard tests in the literature (for example, Imbens and Lemieux (2008)), I first show that there is no manipulation of the assignment variable (as in McCrary (2008)). I then show that individuals in the control and treatment groups are comparable in predetermined characteristics.

First, Figure 2 indicates that the distribution of students in the restricted sample is continuous across the cutoff, meaning that neither the students nor the teachers have manipulated the running variable. Each dot represents the number of students scoring in bins of five PSU-points. The lines correspond to fitted fourth-order polynomials and 95% confidence intervals, estimated for each side of the cutoff separately. Each bin contains approximately 60 students.

Second, each row in Table 2 shows a separate estimation of the model in equation 1, where the dependent variable is some predetermined characteristic, indicated in the first column. Columns (1) to (3) show, respectively, the average for the control group, the jump at the cutoff, and the standard error of the change. For all regressions in these columns, I use a local linear regression over a window of 75 points around the threshold. To describe the robustness of the results, the following three columns use a fourth-order polynomial over the whole PSU support.

For the first specification, the only variable that appears not balanced is an indicator for whether

Figure 2: McCrary test.



Note: Vertical lines at 475 and 550 correspond to the loan and Bicentenario scholarship cutoffs respectively. Each bubble represents the number of students at each bin of 5 PSU-points. Solid lines correspond to a fitted fourth-order polynomial for the running variable at each side of the cutoff.

the father attended college education at a 10% level of significance. However, this imbalance is in line with a type I error. Given the 22 tests shown in the table, about two should be significant at that level of confidence under the null. Additionally, the specification using a fourth-order polynomial over the whole PSU support shows a similar pattern, only one variable is not balanced at the 10%, an indicator for whether the mother attended college, given support to the hypothesis of a valid RD design. [15]

[15] To reinforce this idea I present figures for the raw data figures in the appendix in figure B.2

Table 2: Balance of Covariates

Sample: $Enrolled_{t=1} = 1$ & $Ineligible_{t=1} = 1$ & $q_{t=2} \leqslant 4$						
Specification:	Linear & bw=75			4th Polyn. & whole domain		
Variable	level	jump	se	level	jump	se
Income quintile 1	0.44	-0.05	(0.04)	0.38	0.00	(0.06)
Income quintile 2	0.26	0.05	(0.04)	0.31	0.04	(0.06)
Income quintile 3	0.17	0.01	(0.03)	0.16	-0.01	(0.05)
Income quintile 4	0.13	-0.01	(0.03)	0.15	-0.02	(0.04)
Average quintile	1.99	0.04	(0.09)	2.07	-0.06	(0.13)
Female	0.60	0.00	(0.04)	0.58	0.00	(0.06)
Age at psu	19.6	-0.17	(0.23)	20.0	-0.63	(0.4)
H. School GPA	5.51	-0.01	(0.04)	5.44	0.07	(0.05)
Household size	4.79	-0.20	(0.17)	4.77	-0.22	(0.24)
private H. School	0.03	0.02	(0.02)	0.05	0.02	(0.03)
voucher H. School	0.54	0.00	(0.04)	0.54	-0.02	(0.06)
municipal H. School	0.41	-0.01	(0.04)	0.37	0.02	(0.06)
Health system	2.39	0.19	(0.16)	2.51	0.16	(0.22)
Mother years of educ.	11.6	0.19	(0.3)	11.7	0.31	(0.44)
Father years of educ.	11.5	-0.05	(0.34)	12.1	-0.47	(0.47)
Mother dropout H.S.	0.34	0.02	(0.04)	0.33	0.03	(0.06)
Father dropout H.S.	0.36	0.05	(0.04)	0.33	0.08	(0.06)
Mother college	0.09	0.02	(0.03)	0.09	0.07	(0.04)*
Father college	0.07	0.05	(0.03)*	0.09	0.06	(0.04)
Mother in formal work	0.35	-0.01	(0.04)	0.33	0.01	(0.06)
Father in formal work	0.53	-0.04	(0.04)	0.51	-0.02	(0.06)
Mother is housewife	0.41	0.04	(0.04)	0.40	0.05	(0.06)

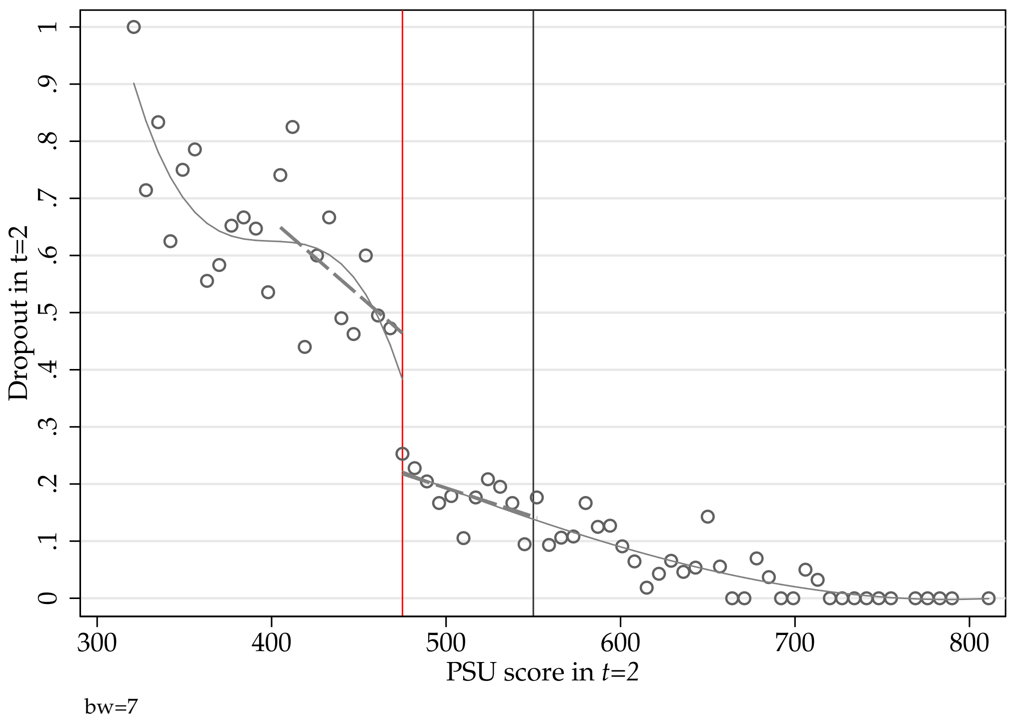
Note: Estimation of equation 1 for the variables in the first column using the sample of analysis. First three columns use a local linear regression (f in equation 1) over a window of 60 PSU-points. Columns (4) to (6) use a fourth-order polynomial f over the whole PSU support. Robust standard error in parentheses. *: $p \leqslant 10\%$, **: $p \leqslant 5\%$, ***: $p \leqslant 1\%$

4.2 Main results

I start by describing the results in a graphical form. Figure 3 shows the dropout rate in the second year ($t = 2$), right after the second PSU attempt. As before, each dot in the figure represents the average dropout rate among students in bins of 7 PSU-points; dashed lines represent fitted values from linear regression at each side of the loan threshold for the bandwidth in the tables, while the solid lines correspond to a quadratic polynomial estimated over the full range of PSU values. The vertical lines represent the loan and the Bicentenario grant cutoffs, at 475 and 550, respectively.

We observe that the dropout rate decreases steadily with ability, starting close to one for very low-ability students and ending with no dropout for the upper extreme of the ability distribution. Importantly, the relationship drops abruptly after students become eligible for loans. Students who score just high enough to qualify for loans at the end of their first year of college reduce their dropout rate approximately from 45% to 25%.

Figure 3: Dropout rate in the second year of college ($t = 2$).



Note: Vertical lines at 475 and 550 correspond to the loan and Bicentenario scholarship cutoffs, respectively. Each dot represents the average dropout rate among students in bins of 7 PSU-points. Solid lines represent a fitted fourth-order polynomial for the running variable at each side of the cutoff.

More formally, Table 3 presents equivalent regression discontinuity regressions for the dropout rate in the second year of college under different specifications. The first three columns use a local linear regression over a bandwidth of 60 points. The first column presents our preferred specification–i.e., the regression in equation 1. The estimates in column (1) confirm the message from Figure 3, showing that students who are barely ineligible for loans in the second year of college are twice as likely as students who are barely eligible for loans to drop out. The dropout rate decreases from 0.47 for the first group to 0.23 for the second.

Table 3: Dropout rate in the second year of college.

	$P(E_{i2}=0 \mid S_0, PSU_{i2})$			$P(V_{i2}=1 \mid \cdot)$	$P(M_{i2}=1 \mid \cdot)$
	(1)	(2)	(3)	(4)	(5)
$1(PSU_2 \geqslant 475)$	-0.24***	-0.22***	-0.22***	-0.15***	-0.085**
	(0.043)	(0.047)	(0.047)	(0.037)	(0.034)
Const.	0.47***	0.42***	0.43***	0.27***	0.20***
	(0.034)	(0.038)	(0.050)	(0.032)	(0.027)
Obs.	1,533	1,288	1,288	1,533	1,533
Controls		x	x		
Years FE			x		

Note: Estimation of equation 1 for dropout rates using the sample of analysis. First three columns use a local linear regression (f in equation 1) over a window of 60 PSU-points. Columns (4) to (5) decompose the effect in column (1) into the destination after dropping out. Robust standard error in parentheses. *: $p \leqslant 10\%$, **: $p \leqslant 5\%$, ***: $p \leqslant 1\%$

The second and third columns explore how robust are the estimates to alternative specifications. Column (2) adds an extensive list of covariates.[16] I find that the estimate is almost the same, confirming the validity of the RD.[17] Column (3) adds year fixed effects to compare students within the same year. In principle, each year is an independent natural experiment, given that we consider only students who have taken the PSU test for the first time after high school graduation and enrolled in college immediately. These observations appear only once in each year. Moreover, the quasi-random nature of the PSU test around the cutoff allows for pooling the students. Effects in column (3) do not vary significantly with respect to column (1), confirming this assumption.

Columns (4) and (5) decompose the result in column (1) as in equation (4). They show that among the 47% of barely ineligible who drop out, 27% went to vocational training, and 20% went to the labor market. Considering that vocational education can be financed through the SGL loans for

[16]Specifically, the covariates are a female indicator, year of birth, mother's and father's education in years, marriage status, affiliation with specific health systems, household size, high school GPA, type of high school, a self-reported category on family income, and an indicator of the student's employment situation.

[17]The constant represents the average dropout rate for the group barely below the cutoff.

all student with high school GPA over 5.3, it is surprising that these dropouts decided not to enroll in the alternative higher education. Similarly, among the 24% of barely eligible dropouts, 12% went to vocational training, and 12% went to the labor market.

5 Mechanisms: Loans Versus Grants

One critical question for policymakers is to determine whether students drop out because of credit market failures or due to changes in the returns to education that are implicit in the financial aid packages. Table 4 attempts to disentangle the mechanisms underlying the reduction in the dropout rate.

To test these mechanisms, I use the Bicentenario Scholarship. Given that crossing the grant threshold allows students to substitute loans for the grant, we can compare students who finance the same amount with a different aid tool. If students who become eligible for the scholarship reduce their dropout rate as in 475, it would be an indication that the effect of the loans is driven by perceived subsidies that are present in the loans.

Table 4 follows the same structure as in Table 3.[18] We observe that the dropout rate is the same for students across this threshold, which means that students are not sensitive to tuition costs, and access to financial markets was the primary determinant of the change found at the loan cutoff. The effect is robust to the inclusion of covariates (column (2)) and year fixed effects (column (3)); and the students who drop out (14 percentage points) have very similar responses, moving to vocational training and the labor market (columns (4) and (5)).

[18]The number of observations is lower because the two poorest income quintiles are eligible.

Table 4: Dropout rate in the second year, around the Bicentenario scholarship.

	$\mathrm{P}(E_{i2}=0\|S_0, PSU_{i2})$			$\mathrm{P}(V_{i2}=1\|\cdot)$	$\mathrm{P}(M_{i2}=1\|\cdot)$
	(1)	(2)	(3)	(4)	(5)
$1(PSU_2 \geqslant 550)$	0.032	0.031	0.038	0.023	0.0095
	(0.054)	(0.058)	(0.058)	(0.037)	(0.042)
Const.	0.14^{***}	15.9	4.77	0.062^{**}	0.081^{***}
	(0.035)	(11.6)	(16.8)	(0.025)	(0.028)
Obs.	837	711	711	837	837
Controls		x	x		
Years FE			x		

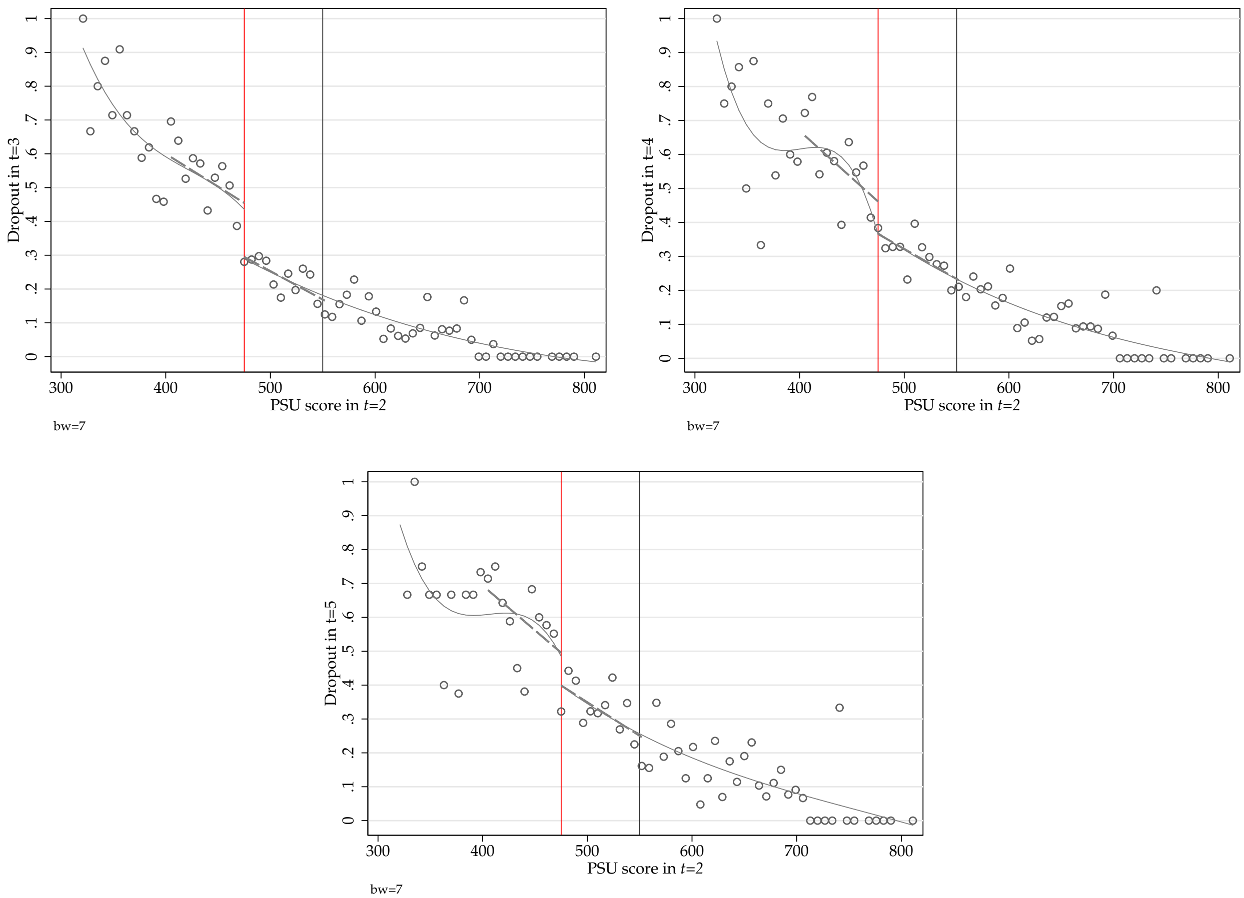
Note: Estimation of equation 1 for dropout rates around the Bicentenario Scholarship using the conditional sample described in the text. First three columns use a local linear regression and a window of 60 PSU-points. Columns (4) to (6) uses a fourth-order polynomial f over the whole PSU domain. Robust standard error in parentheses. *: $p \leqslant 10\%$, **: $p \leqslant 5\%$, ***: $p \leqslant 1\%$

6 Longer-Run Effects

Policies preventing dropout in the second year would not affect the attainment gap if students affected by the policy drop out later, before graduation. In this section, I explore the dynamic effects of this policy, examining dropout in the third, fourth, and fifth years.

I start with a graphical analysis. Figure 4 presents the dropout rate for the following years. Here, the dropout rate is defined as an indicator of no enrollment in the designated year. The analysis in these figures restricts the sample to the cohorts that may achieve the year of college under study. For example, when studying the dropout rate at the fifth year, I use cohorts 2007 to 2011 because those who enrolled in 2011 achieve the fifth year of college by the year 2016, which is the last year of enrollment data available for this paper.

Figure 4: Dropout rate. Population restricted to cohort 2007-2010



Note: Vertical lines at 475 and 550 correspond to the loan and Bicentenario scholarship cutoffs, respectively. Each bubble represents the average dropout rate among students in bins of 7 PSU-points. Solid lines represent a fitted fourth-order polynomial for the running variable at each side of the cutoff.

These figures correspond to the reduced-form estimation. For those barely above the cutoff, the dropout rates increase steadily over the years, starting at 30% and, achieving 40% in the fifth year. Remarkably, the change in the dropout rate at the cutoff is stable over time (from 10 to 15 percentage points).

Tables 5 and 6 present the corresponding fuzzy regression discontinuity, as in equations (2) and (3), for the third and fourth, respectively.[19] These tables address the fact that some students self-select into treatment and, therefore, I instrument eligibility for loans in any period $t > 2$ by the indicator for scoring at least the cutoff in $t = 2$.

[19]Results for the fifth year are left in Table D.1 in the appendix

Table 5: Dropout in the third year of college

	Overall	Time decomposition		Decomposition by destination			
	$p_{3,.,0}$	$p_{3,1,0}$	$p_{3,0,0}$	$p_{3,1,v}$	$p_{3,1,m}$	$p_{3,0,v}$	$p_{3,0,m}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. 2SLS:							
Eligile for Loans	-0.18***	0.022	-0.21***	-0.0062	0.029	-0.20***	-0.0056
	(0.054)	(0.031)	(0.050)	(0.017)	(0.027)	(0.046)	(0.030)
Const.	0.47***	0.070***	0.41***	0.029**	0.040**	0.32***	0.080***
	(0.042)	(0.021)	(0.041)	(0.012)	(0.018)	(0.039)	(0.024)
Obs.	1,315	1,315	1,315	1,315	1,315	1,315	1,315
B. Reduced form:							
$1(PSU_2 \geqslant 475)$	-0.16***	0.020	-0.18***	-0.0056	0.026	-0.18***	-0.0050
	(0.049)	(0.028)	(0.045)	(0.015)	(0.024)	(0.041)	(0.027)
Const.	0.46***	0.072***	0.38***	0.029***	0.043***	0.30***	0.080***
	(0.038)	(0.019)	(0.037)	(0.011)	(0.016)	(0.036)	(0.021)
Obs.	1,315	1,315	1,315	1,315	1,315	1,315	1,315

Note: Estimation of equation 1 for dropout rates using the sample of analysis. First three columns use a local linear regression (f in equation 1) over a window of 60 PSU-points. Columns (4) to (6) uses a fourth-order polynomial f over the whole PSU support. Robust standard error in parentheses. *: $p \leqslant 10\%$, **: $p \leqslant 5\%$, ***: $p \leqslant 1\%$

Panel A in Table 5 shows the 2SLS estimation of the dropout rate and its decompositions (temporal and by destination), and Panel B shows the corresponding reduced forms. Column (1) shows that the dropout rate among the barely ineligible is the same as in $t = 2$, 47%, but now, occurs as a combination of students who dropped after two years of enrollment and student who came back to college in the third after missing a year. However, access to loans leads to a difference in the dropout rate of 18 percentage points, at the cutoff.

Columns (2) and (3) show the temporal decomposition, as in equation (5). Column (2) shows the proportion of students who were enrolled in two consecutive years but did not enroll in the third, while Column (3) presents the share of students who did not enroll in the second year and did not come back. These columns indicate that most of the dropout happened in the second year, and after that year, the difference in the dropout rate is statistically zero.

Columns (4) to (7) describe the alternative chosen by the students who dropped out. Students who enrolled in two consecutive years but failed to enroll in the third (columns (4) and (5)) split their decision between going to vocational school or the labor market at very similar rates, about 3 percentage points chose either alternative. Finally, columns (6) and (7) illustrate the destination of students who dropped out in the second year and never came back. They show that among the barely ineligible, 78% (32/41) of the students chose vocational education. Among the eligibles, the split is more similar, at 11pp and 7pp for vocational school and the labor market, respectively.[20]

Table 6 shows 2SLS estimates for dropping out in the fourth year of college and its decomposition over time and by destination after college.[21] Panel A shows that the dropout rate for the barely ineligible is 50%, and decreases by 13 percentage points at the cutoff. Columns (2) and (3) compare the behavior among students, depending on whether or not they were enrolled in the second year. Column (2) shows that the dropout rate did not change for those who survived to the second year. For both groups, the dropout rate was approximately 13-15%. Column (3), shows, instead, that the bulk of the 52% dropout rate for the ineligible in $t = 4$ corresponds to students who dropped out in the second year, while only 17% of those who became eligible in $t = 2$ dropped out in the second year and never came back. In columns (4) to (7), we see the destination of eligible and ineligible students. For those who survived to the second year, there are no statistical differences in the outside option. They seem to choose equally to go to vocational school or the labor market. In columns (6) and (7), the destinations differ substantially. Among dropouts in the second year who never came back, 4% of the barely eligible chose the labor market and 13% vocational programs. Among the barely ineligible, 15% went to the labor market and 24% to vocational programs.

Panel B decomposes the effect over time, into $p_{4,0,0,0}$, the part of the non-enrollment gap in $t = 4$ that is attributable to people who dropped out in $t = 2$ and then never returned. The RD is decomposed into $p_{4,0,1,0}$ the part due to people who dropped in $t = 2$, came back in $t = 3$, but dropped out again in the fourth year. The RD for $p_{4,1,0,0}$ corresponds to individuals who enrolled in

[20]Corresponding reduced-form estimates are shown in Appendix D.2.
[21]Formal estimations of the reduced forms for $t = 4$ and 5 are given in Appendix D.2

the second year but then dropped out in year 3 and never came back. Finally, the RD for $p_{4,1,1,0}$ is the part attributable to people who stayed enrolled for the second and third but then dropped out in the fourth year.

We observe that most of the action occurred in the second year. Among ineligible students, 37% dropped out and never came back; 2% enrolled again in the third year but failed to enroll in the fourth; and about 6% dropped out in the fourth year after being continuously enrolled. For those who became eligible in $t = 2$, the dropout rate in that year was 20 percentage points lower and no different from the rate for ineligible in all other years.

Table 6: Dropout rate in the fourth year of college

[A.] Decomposition by destination. 2SLS:

	$p_{4,.,.,0}$	$p_{4,1,.,v}$	$p_{4,1,.,m}$	$p_{4,0,.,v}$	$p_{4,0,.,m}$
	(1)	(2)	(3)	(4)	(5)
Eligile for Loans	-0.13**	-0.02	0.05	-0.11**	-0.05
	(0.060)	(0.028)	(0.035)	(0.049)	(0.035)
Const.	0.50***	0.07***	0.07***	0.24***	0.11***
	(0.046)	(0.022)	(0.023)	(0.041)	(0.029)
Obs.	1098	1098	1098	1098	1098

[B.] Temporal Decomposition. 2SLS:

	$p_{4,1,.,0}$	$p_{4,0,.,0}$	$p_{4,0,0,0}$	$p_{4,0,1,0}$	$p_{4,1,0,0}$	$p_{4,1,1,0}$
	(1)	(2)	(3)	(4)	(5)	(6)
Eligible for Loans	0.03	-0.16***	-0.16***	-0.00	-0.01	0.04
	(0.043)	(0.054)	(0.054)	(0.013)	(0.030)	(0.034)
Const.	0.14***	0.36***	0.35***	0.01	0.06***	0.08***
	(0.031)	(0.044)	(0.044)	(0.011)	(0.022)	(0.024)
Obs.	1098	1098	1098	1098	1098	1098

Note: Estimation of equation 1 for dropout rates using the sample of analysis. First three columns use a local linear regression (f in equation 1) over a window of 60 PSU-points. Columns (4) to (6) uses a fourth-order polynomial f over the whole PSU support. Robust standard error in parentheses. *: $p \leqslant 10\%$, **: $p \leqslant 5\%$, ***: $p \leqslant 1\%$

Similar patterns appear for the fifth year: most students dropped out in the second year, and they were equally likely to move to vocational institutions or to enter the labor market directly. For ease of exposition, I present the results in Table D.1 in the Appendix.

6.1 Effects by Family Income

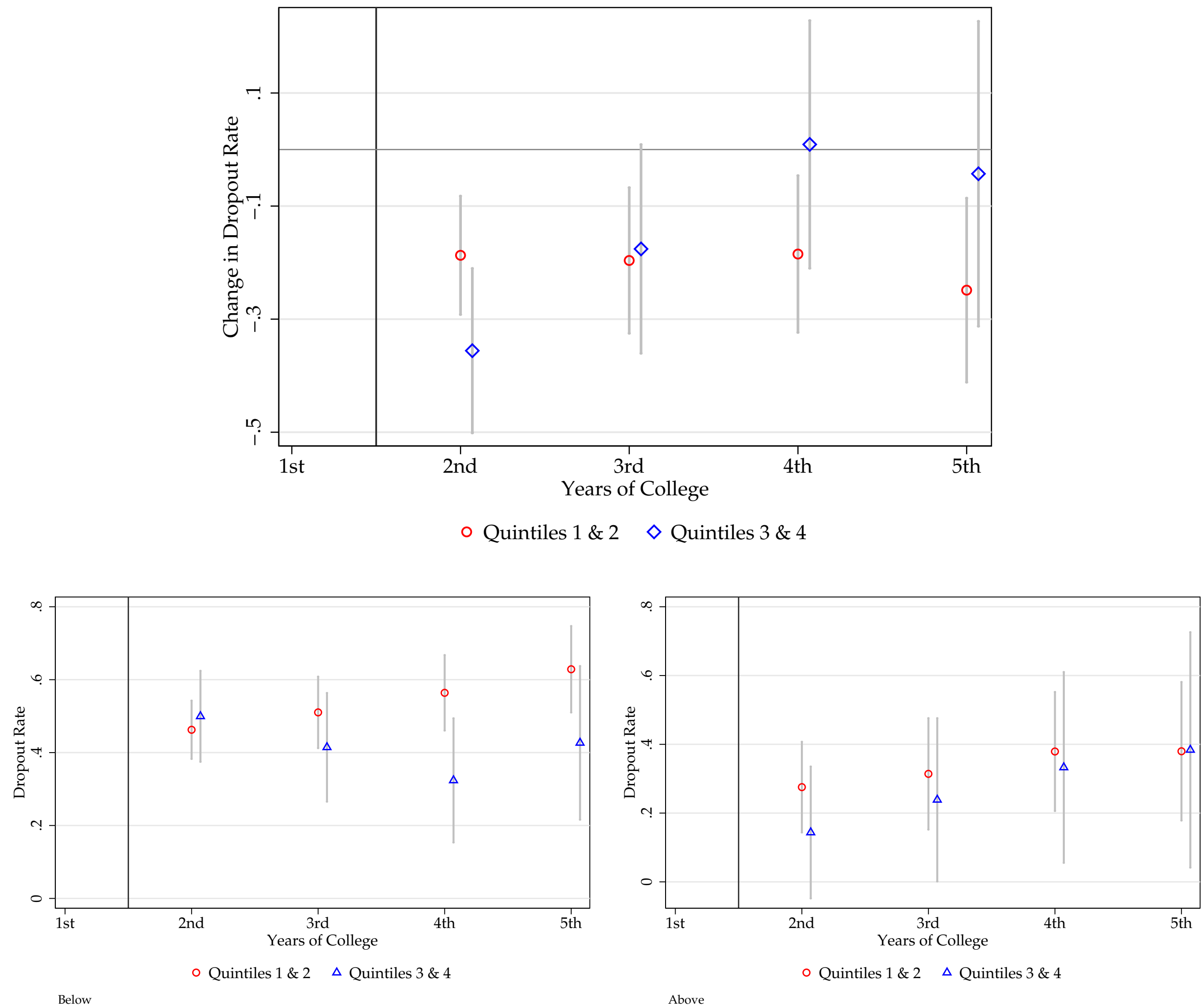
Now, we turn to the analysis of the gap in education attainment by family income. Given the small number of observations, I pooled quintiles into two groups: quintiles one and two (the poorest from now on) form one group, and quintiles three and four together (the richer from now on) form the other. Figure E.1 in the appendix shows the RDs for each group in the four years of college that I study. To summarize, I plot the estimates at the cutoff for each income group in Figure 5. The figure at the top shows the estimated jump at each level, and the figures at the bottom show the estimated levels of the dropout rate.

For the poorest, the effect on the dropout rate at the cutoff is close to 20 percentage points in every year, as shown in Figure 5 (red circles). In contrast, for the richer, the effect is slightly higher for the first two years and then becomes very close to zero (blue triangles). The figures at the bottom show the dropout rates for both groups (for students with access to loans in the right, and students without access on the left). Ineligible students have a much higher dropout rate from the start. For the poor, the level grows steadily over time while, for the richer, the level decreases and, by the fourth year become very similar to the dropout rate for eligible students.

In contrast, for the students with access to loans, the figure on the right shows that the evolution of the dropout rate is very similar for the poorest and the richer. At the fifth year of college, the dropout rates are practically the same, 40%. Consequently, eligibility for loans allows the poorest income quintiles to have the same dropout rate as the richer income quintile, thus closing the gap.

Table 7 shows the formal estimation. Each pair of columns considers a different college year for the two groups. In the table, we find the same patterns observed in the figures. For example, we observe the stability of the estimation for the poorest: 20 percentile points each year. For the richer, in contrast, in the fourth year, the jump at the cutoff is 0.004.

Figure 5: Dropout rate in the second year of college ($t = 2$).



Note: The top figure depicts the estimated effect at the cutoff by income group and its corresponding 95% confidence interval. On the x-axis is the year of college for the estimation. The bottom shows the level of the dropout rate, for ineligible students at the left and the eligible at the right.

6.2 Effects by gender

The effects of loan eligibility on the dropout rate may vary by gender, which could be an indication of gender preferences or barriers to entering the labor market that differ by gender. For example, it may be harder for females to find part-time jobs to fund college, reflecting the lower female labor market participation. In a country with patriarchal values, parents may prefer to finance the

Table 7: Dropout rate from the second to the fifth year of college by income. 2SLS estimates.

	$t=2$		$t=3$		$t=4$		$t=5$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Decomposition by income quintile:								
	q_1q_2	q_3q_4	q_1q_2	q_3q_4	q_1q_2	q_3q_4	q_1q_2	q_3q_4
$1(PSU_2 \geqslant 475)$	-0.20***	-0.34***						
	(0.053)	(0.074)						
Eligile for Loans			-0.19***	-0.16*	-0.18***	0.0083	-0.25***	-0.045
			(0.066)	(0.094)	(0.071)	(0.11)	(0.083)	(0.14)
Const.	0.46***	0.48***	0.51***	0.40***	0.56***	0.32***	0.63***	0.43***
	(0.041)	(0.064)	(0.050)	(0.075)	(0.053)	(0.087)	(0.060)	(0.11)
Obs.	1,003	535	856	459	734	368	588	281

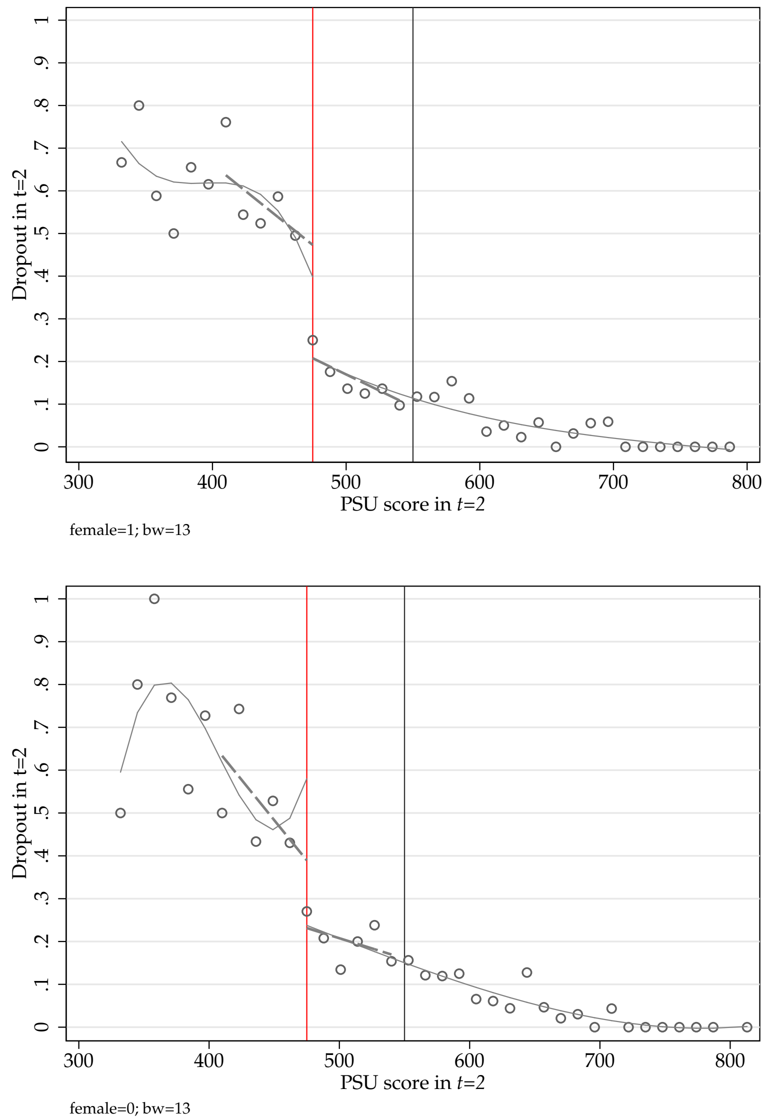
Note: Bandwidth equal to 80 points.

education of boys, given that men need to be perceived as financially independent to succeed in the marriage market. As a consequence, access to credit may be more important for women.

Figure 6 and Table 8 document the differences by gender. Figure 6 shows that the effect of credit access on the dropout rate in the second year is stronger for females. Males and females have similar dropout rates below the cutoff (about 45%), but it decreases by 25 percentage points for women vs. 20 percentage points for men. However, the difference is not statistically significant.

Table 8 gives the precise RD estimation and expands the analysis to subsequent years and for both decompositions studied earlier. Panel A shows that the chosen outside option differs greatly by gender. In the second year of college, women chose vocational training as their most preferred alternative, while men preferred entering the labor market. This preference was the same for ineligible and eligible students indistinctly. For example, among dropout students, women chose vocational education 67% (0.33/0.49) of the time and men 42% (0.18/0.43)

Figure 6: Dropout rate by gender



Note: Vertical lines at 475 and 550 correspond to the loan and scholarship cutoffs, respectively. Each bubble represents the average dropout rate among students in bins of 13 PSU-points. Solid lines represent a fitted fourth-order polynomial for the running variable at each side of the cutoff.

Loans equalize the substitution patterns. Only 14% of females migrated to vocational programs after getting access to loans (0.33 -0.19). These numbers imply that for every non-constrained women switching to vocational programs, almost 2.4 (0.14/0.33) did the same in the absence of loans. The equivalent for men was 2 (0.09/.18). Access to loans seems not to have affected (statistically) the decision to move to the labor market. About 8.5% of women and 12% of men migrated to the labor market.

Table 8: Dropout rates by gender.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Dropout in second by gender:							
	Female			Male			
	$p_{2,0}$	$p_{2,v}$	$p_{2,m}$	$p_{2,0}$	$p_{2,v}$	$p_{2,m}$	
$1(PSU_2 \geqslant 475)$	-0.26***	-0.19***	-0.065	-0.21***	-0.087	-0.13**	
	(0.056)	(0.050)	(0.040)	(0.068)	(0.055)	(0.056)	
Const.	0.49***	0.33***	0.15***	0.43***	0.18***	0.25***	
	(0.044)	(0.042)	(0.033)	(0.055)	(0.048)	(0.046)	
Obs.	889	889	889	649	649	649	
B. Dropout in third: Female. 2SLS:							
	$p_{3,.,0}$	$p_{3,1,0}$	$p_{3,0,0}$	$p_{3,1,v}$	$p_{3,1,m}$	$p_{3,0,v}$	$p_{3,0,m}$
Eligile for Loans	-0.21***	0.017	-0.22***	-0.025	0.042	-0.24***	0.018
	(0.070)	(0.042)	(0.065)	(0.021)	(0.037)	(0.061)	(0.033)
Const.	0.50***	0.071***	0.43***	0.037**	0.035	0.38***	0.044*
	(0.054)	(0.027)	(0.053)	(0.017)	(0.021)	(0.051)	(0.025)
Obs.	743	743	743	743	743	743	743
C. Dropout in third: Male. 2SLS:							
	$p_{3,.,0}$	$p_{3,1,0}$	$p_{3,0,0}$	$p_{3,1,v}$	$p_{3,1,m}$	$p_{3,0,v}$	$p_{3,0,m}$
Eligile for Loans	-0.14*	0.030	-0.17**	0.020	0.0097	-0.13*	-0.041
	(0.084)	(0.048)	(0.080)	(0.026)	(0.041)	(0.069)	(0.056)
Const.	0.43***	0.066*	0.37***	0.019	0.048	0.23***	0.13***
	(0.067)	(0.034)	(0.066)	(0.016)	(0.031)	(0.060)	(0.046)
Obs.	572	572	572	572	572	572	572

Note: Estimation of equation 1. All columns use a local linear regression (f in equation 1) over a window of 60 PSU-points. In Panel A, the first three columns are the dropout rate in the second year and its decomposition between vocational and labor market decisions for females. Columns (4) to (6) for males. Panel B shows the dropout rate for women in the third year of college and its decomposition. Panel C for men. Robust standard error in parentheses. *: $p \leqslant 10\%$, **: $p \leqslant 5\%$, ***: $p \leqslant 1\%$

For males, the rate of migration to vocational education was 10 percentage points, compared to 12 percentage points for women.[22] Moreover, the rate of substitution with the labor market was 15

[22] I cannot reject the hypothesis that these rates are the same.

percentage points, compared to 10 percentage points for females.

Panels B and C show 2SLS estimations for the dropout rate in the third year of college. Column (1) in panels B and C shows that the dropout rate for eligible and ineligible students of both genders is quite similar (50 vs. 43 percentage points for the ineligible – female and males respectively– and 29 and 27 percentage points for eligibles). Columns (2) and (3) show the temporal decomposition and columns (4) to (6) the decomposition by destination for the dropout rate in $t = 3$. Column (3) confirms the previous discussion about the importance of the second year. Most of the non-enrollment rate in the third year occurs right after the second PSU attempt: among the ineligible, 88% (.413/.472) for females vs. 82% (.354/.434) for males.

The chosen outside option differs by gender as well. Among ineligible females, virtually all opted for vocational programs when they dropped out in the second year, (96%=0.397/0.413), while ineligible males chose the vocational and labor market options equally. Eligible males and females went evenly into vocational education and the labor market; roughly 10% chose each.

This evidence shows that females strongly prefer an educational option rather than the labor market. Females may have stronger preferences for education because of the broader educational returns in the marriage market or because of higher barriers to entry in the labor market.[23]

7 Conclusion

In this paper, I show the effects of access to loans on the college dropout decisions of inframarginal students in Chile. I use variation from an eligibility cutoff for college loans to perform an RD analysis.

Using arguably exogenous variation in loan access that affects only the decision to continue college but not enrollment, I find that students without access to credit in the second year are twice as likely to drop out than those who have access. The dropout rate for students with no access is

[23]The labor market participation rate for women between ages 18 and 23 is 30%, compared to 55% for men, as indicated by CASEN (2011) (CASEN is the most important household survey in Chile).

47% in the second year of college and, is reduced to 23% for those who benefit from crossing the threshold. The remains roughly unaltered in the following years up to the fifth years of college. Importantly, the analysis of the Bicentenario Scholarship allows us to conclude that access to loans drives the effects, and not the presence of underlying subsidies.

The paper also shows that the earlier years are the most crucial when students are still learning about their true ability type. Interestingly, not all ineligible students move to vocational education where financing is available, and some move directly to the labor market, indicating a perceived difference in the returns to different types of education. The effects are stronger to women, suggesting small differences in the taste for education and the barriers to enter the labor market. However, statistical differences are not detectable.

The evidence shows that loan programs are useful to reduce the attainment gap between students from rich and poor backgrounds. Using the income classification made by the tax authority, we can reconstruct the educational attainment gap by family income around the eligibility cutoff. I find that low-income students are 20 percentage points more likely to drop out when they do not have access to loans. In contrast, the gap disappears when students become eligible for loans. In this context, college is relatively expensive; thus, the absence of policies to facilitate access to college may have a significant contribution to intergenerational inequality.

The context of this paper is different from most of the literature. Regarding the evidence from developed countries, most of the evidence relies on the study of grants, and the context is relative more generous. In the US and many other developed countries, policies rest in an extensive and well-functioning aid system, where changes in aid affect students less. This paper can shed light about the consequences of loans when other types of aid are scarce.

The evidence presented in the paper is also relevant for policymakers. Loans are one of the most common forms of financial aid, especially in developing countries. [24] The design of financial aid policies, in similar contexts, may increase efficiency and equity.

[24] According to the International Comparative Higher Education and Finance Project, from the State University of New York at Buffalo, from the 47 countries described 41 have some form of a loan.

References

Angrist, J., D. Autor, S. Hudson, and A. Pallais (2016): “Evaluating Post-Secondary Aid: Enrollment, Persistence, and Projected Completion Effects,” Working Paper 23015, National Bureau of Economic Research.

Angrist, J., D. Lang, and P. Oreopoulos (2009): “Incentives and Services for College Achievement: Evidence from a Randomized Trial,” *American Economic Journal: Applied Economics*, 1, 136–163.

Barrios Fernández, A. (2019): “Should I Stay or Should I Go? Neighbors’ Effects on University Enrollment,” SSRN Scholarly Paper ID 3278251, Social Science Research Network, Rochester, NY.

Bettinger, E. (2004): “How Financial Aid Affects Persistence,” Working Paper 10242, National Bureau of Economic Research.

Brock, T. and L. Richburg-Hayes (2006): *Paying for Persistence. Early Results of a Louisiana Scholarship Program for Low-Income Parents Attending Community College*, MDRC, 16 East 34th Street, 19th Floor, New York, NY 10016-4326.

Bucarey, A., D. Contreras, and P. Munoz (2018): “Labor Market Returns to Student Loans,” SSRN Scholarly Paper ID 3173823, Social Science Research Network, Rochester, NY.

Castleman, B. L. and B. T. Long (2016): “Looking beyond Enrollment: The Causal Effect of Need-Based Grants on College Access, Persistence, and Graduation,” *Journal of Labor Economics*, 34, 1023–1073.

Cellini, S. R. (2009): “Crowded Colleges and College Crowd-Out: The Impact of Public Subsidies on the Two-Year College Market,” *American Economic Journal: Economic Policy*, 1, 1–30.

Cellini, S. R., R. Darolia, and L. J. Turner (Forthcoming): “Where Do Students Go when For-Profit Colleges Lose Federal Aid?” *American Economic Journal: Economic Policy.*

Chen, R. (2008): "Financial Aid and Student Dropout in Higher Education: A Heterogeneous Research Approach," in *Higher Education*, ed. by J. C. Smart, Springer Netherlands, no. 23 in Handbook of Theory and Research, 209–239.

Cohodes, S. R. and J. S. Goodman (2014): "Merit Aid, College Quality, and College Completion: Massachusetts' Adams Scholarship as an In-Kind Subsidy," *American Economic Journal: Applied Economics*, 6, 251–285.

Denning, J. T. (2018): "Born Under a Lucky Star: Financial Aid, College Completion, Labor Supply, and Credit Constraints," *Journal of Human Resources*.

Denning, J. T., B. M. Marx, and L. J. Turner (Forthcoming): "ProPelled: The Effects of Grants on Graduation, Earnings, and Welfare," *American Economic Journal: Applied Economics*.

DesJardins, S. L., D. A. Ahlburg, and B. P. McCall (2002): "Simulating the Longitudinal Effects of Changes in Financial Aid on Student Departure from College," *The Journal of Human Resources*, 37, 653–679.

Dynarski, S. and J. Scott-Clayton (2013): "Financial Aid Policy: Lessons from Research," Working Paper 18710, National Bureau of Economic Research.

Dynarski, S. M. (2003): "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion," *The American Economic Review*, 93, 279–288.

Field, E. (2009): "Educational Debt Burden and Career Choice: Evidence from a Financial Aid Experiment at NYU Law School," *American Economic Journal: Applied Economics*, 1, 1–21.

Garibaldi, P., F. Giavazzi, A. Ichino, and E. Rettore (2011): "College Cost and Time to Complete a Degree: Evidence from Tuition Discontinuities," *The Review of Economics and Statistics*, 94, 699–711.

Goldrick-Rab, S., D. Harris, R. Kelchen, and J. Benson (2012): "Need-Based Financial Aid and College Persistence Experimental Evidence from Wisconsin," SSRN Scholarly Paper ID 1887826, Social Science Research Network, Rochester, NY.

Goldrick-Rab, S., D. N. Harris, and P. A. Trostel (2009): "Why Financial Aid Matters (or Does Not) for College Success: Toward a New Interdisciplinary Perspective," in *Higher Education: Handbook of Theory and Research*, ed. by J. C. Smart, Springer Netherlands, no. 24 in Higher Education: Handbook of Theory and Research, 1–45.

Hossler, D., M. Ziskin, J. P. K. Gross, S. Kim, and O. Cekic (2009): "Student Aid and Its Role in Encouraging Persistence," in *Higher Education: Handbook of Theory and Research*, ed. by J. C. Smart, Springer Netherlands, no. 24 in Higher Education: Handbook of Theory and Research, 389–425.

Imbens, G. W. and T. Lemieux (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142, 615–635.

Lee, D. S. (2008): "Randomized experiments from non-random selection in U.S. House elections," *Journal of Econometrics*, 142, 675–697.

McCrary, J. (2008): "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142, 698–714.

Montoya, A. M., C. Noton, and A. Solis (2017): "Returns to Higher Education: Vocational Education vs College," SSRN Scholarly Paper ID 3106354, Social Science Research Network, Rochester, NY.

Murphy, R. and G. Wyness (2016): "Testing Means-Tested Aid," SSRN Scholarly Paper ID 2866923, Social Science Research Network, Rochester, NY.

Rothstein, J. and C. E. Rouse (2011): "Constrained after college: Student loans and early-career occupational choices," *Journal of Public Economics*, 95, 149–163.

Santelices, M. V., X. CatalÃ¡n, D. Kruger, and C. Horn (2015): “Determinants of persistence and the role of financial aid: lessons from Chile,” *Higher Education*, 1–20.

Scott-Clayton, J. (2011): “On Money and Motivation A Quasi-Experimental Analysis of Financial Incentives for College Achievement,” *Journal of Human Resources*, 46, 614–646.

Singell, Jr., L. D. (2004): “Come and stay a while: does financial aid effect retention conditioned on enrollment at a large public university?” *Economics of Education Review*, 23, 459–471.

Solis, A. (2013): “Does higher education cause political participation? Evidence from a regression discontinuity design,” Tech. Rep. 2013:13, Working Paper, Department of Economics, Uppsala University.

——— (2017): “Credit Access and College Enrollment,” *Journal of Political Economy*, 125, 562–622.

Stinebrickner, R. and T. Stinebrickner (2008): “The Effect of Credit Constraints on the College Drop-Out Decision: A Direct Approach Using a New Panel Study,” *The American Economic Review*, 98, 2163–2184.

World Bank (2011): “Chile’s State-Guaranteed Student Loan Program (CAE),” Tech. Rep. 70209, The World Bank.

A Appendix: Decomposition and notation

The dropout probability in the year s is denoted as $P(E_{2i} = j_2, ..., E_{si} = j_s|\cdot) = p_{s,j_2,...,j_s}$, where j_2 indicates the alternative chosen in $t = 2$, and j_s follows. Alternatives j can take the value 0, 1, v or m. For example, when $j = 1$ ($E_{si} = j = 1$), implies the student i enrolled in college in time s. In similar way $j = 0$ means the student drop out from college, $j = v$ means that the students choose to enroll in a vocational program, and $j = m$, implies that the students did not enroll in any for of tertiary education, which we interpret as entering the labor market:

The rate of dropout in the second and third year are given i the main text. Here we focus in the fourth and fifth period. The probability $P(E_{i4} = 0|\cdot)$ ca be decomposed in several ways. Here, for example, we show how these probability can be explained the outcomes in two periods:

$$P(E_{i4} = 0|\cdot) = P(E_{i2} = 1, E_{i4} = 0|\cdot) + P(E_{i2} = 0, E_{i4} = 0|\cdot) \tag{8}$$
$$p_{4,.,.,0} = p_{4,1,.,0} + p_{4,1,.,0}$$

Also, we could have a four-part temporal decomposition:

$$\begin{aligned} P(E_{i4} = 0|\cdot) = {} & P(E_{i2} = 0, E_{i3} = 0, E_{i4} = 0|\cdot) + P(E_{i2} = 1, E_{i3} = 0, E_{i4} = 0|\cdot) \\ & + P(E_{i2} = 0, E_{i3} = 1, E_{i4} = 0|\cdot) + P(E_{i2} = 1, E_{i3} = 1, E_{i4} = 0|\cdot) \end{aligned} \tag{9}$$
$$p_{4,.,.,0} = p_{4,0,0,0} + p_{4,0,1,0} + p_{4,1,0,0} + p_{4,1,1,0}$$

Again, I can estimate RD models for all four terms on the right-hand side. The RD in $P(E_{i2} = 0, E_{i3} = 0, E_{i4} = 0|S; T_{2i})$ is the part of the non-enrollment gap in period $t = 4$ that is attributable to people who dropped out in year $t = 2$ and then never returned. The RD in $P(E_{i2} = 1, E_{i3} = 0, E_{i4} = 0|S; T_{21})$ is the part attributable to people who stayed for the second year but then dropped out in the third year and never came back. And, finally the RD in $P(E_{i2} = 1, E_{i3} = 1, E_{i4} = 0|S; T_{21})$ is the part attributable to people who stayed enrolled for the second and third years but then dropped out in the fourth year.

Similarly, I can decompose $P(E_{i4} = 0|\cdot)$ depending on the alternative activity chosen:

$$
\begin{aligned}
P(E_{i4}=0|\cdot) &= P(E_{i2}=1, E_{i4}=0|\cdot) + P(E_{i2}=0, E_{i4}=0|\cdot) \\
&= P(E_{i2}=1, V_{i4}=1|\cdot) + P(E_{i2}=1, M_{i4}=1|\cdot) \\
&+ P(E_{i2}=0, V_{i4}=1|\cdot) + P(E_{i2}=0, M_{i4}=1|\cdot) \\
p_{4,.,.,0} &= p_{4,1,.,v} + p_{4,1,.,m} + p_{4,0,.,v} + p_{4,0,.,m}
\end{aligned} \tag{10}
$$

For the probability of no enrollment in the fifth year we show the following temporal decomposition:

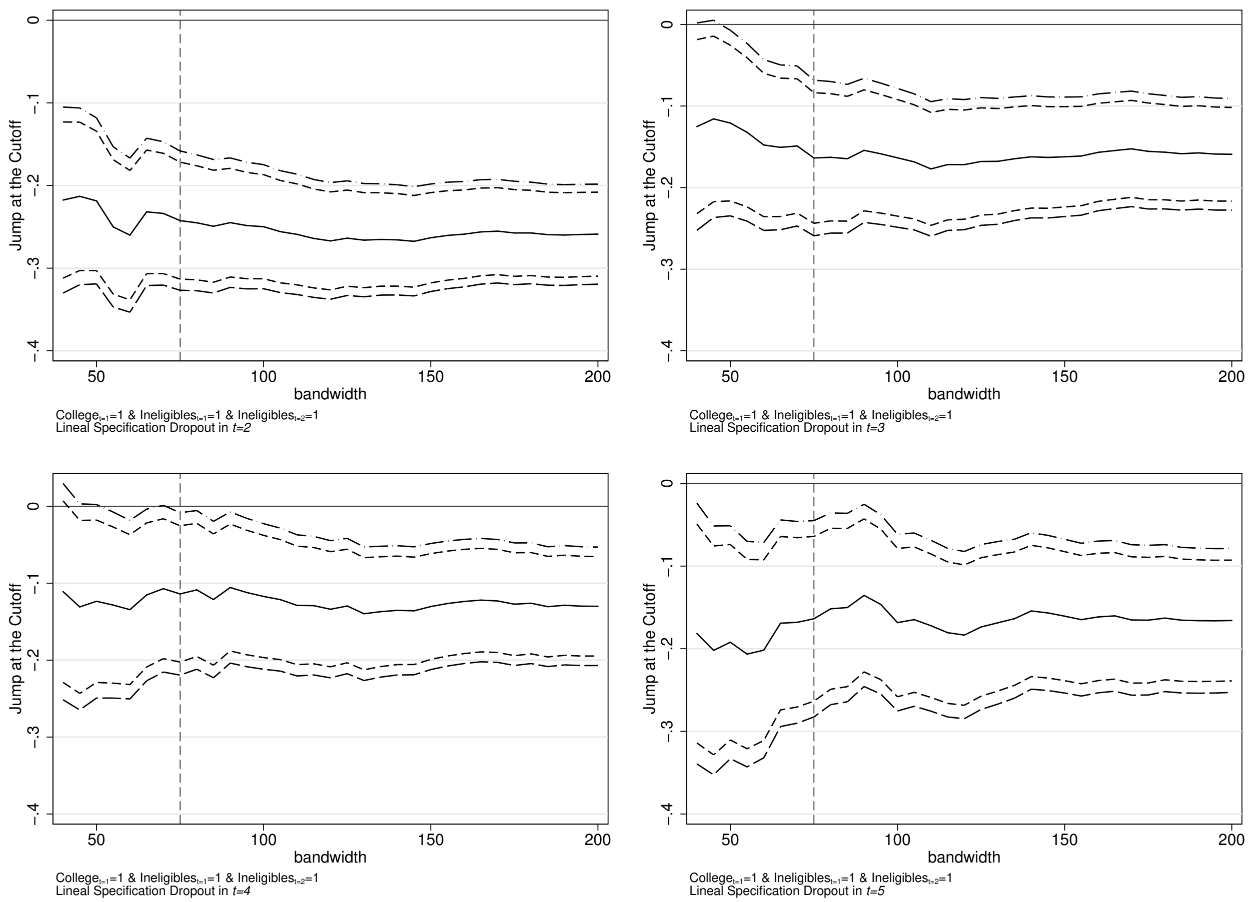
$$
\begin{aligned}
P(E_{i5}=0|\cdot) &= P(E_{i2}=1, E_{i5}=0|\cdot) + P(E_{i2}=0, E_{i5}=0|\cdot) \\
&= P(E_{i2}=1, V_{i5}=1|\cdot) + P(E_{i2}=1, M_{i5}=1|\cdot) \\
&+ P(E_{i2}=0, V_{i5}=1|\cdot) + P(E_{i2}=0, M_{i5}=1|\cdot) \\
p_{5,.,.,.,0} &= p_{5,1,.,.,v} + p_{5,1,.,.,m} + p_{5,0,.,.,v} + p_{5,0,.,.,m}
\end{aligned} \tag{11}
$$

where I care only to what happened in $t = 2$ and the period of analysis $t = s$, $s = 4, 5$.

B Appendix. Additional validation for the RD design

B.1 Sensitivity to Bandwidth

Figure B.1: Sensitivity for Dropout Rate. Jump for different bandwidths



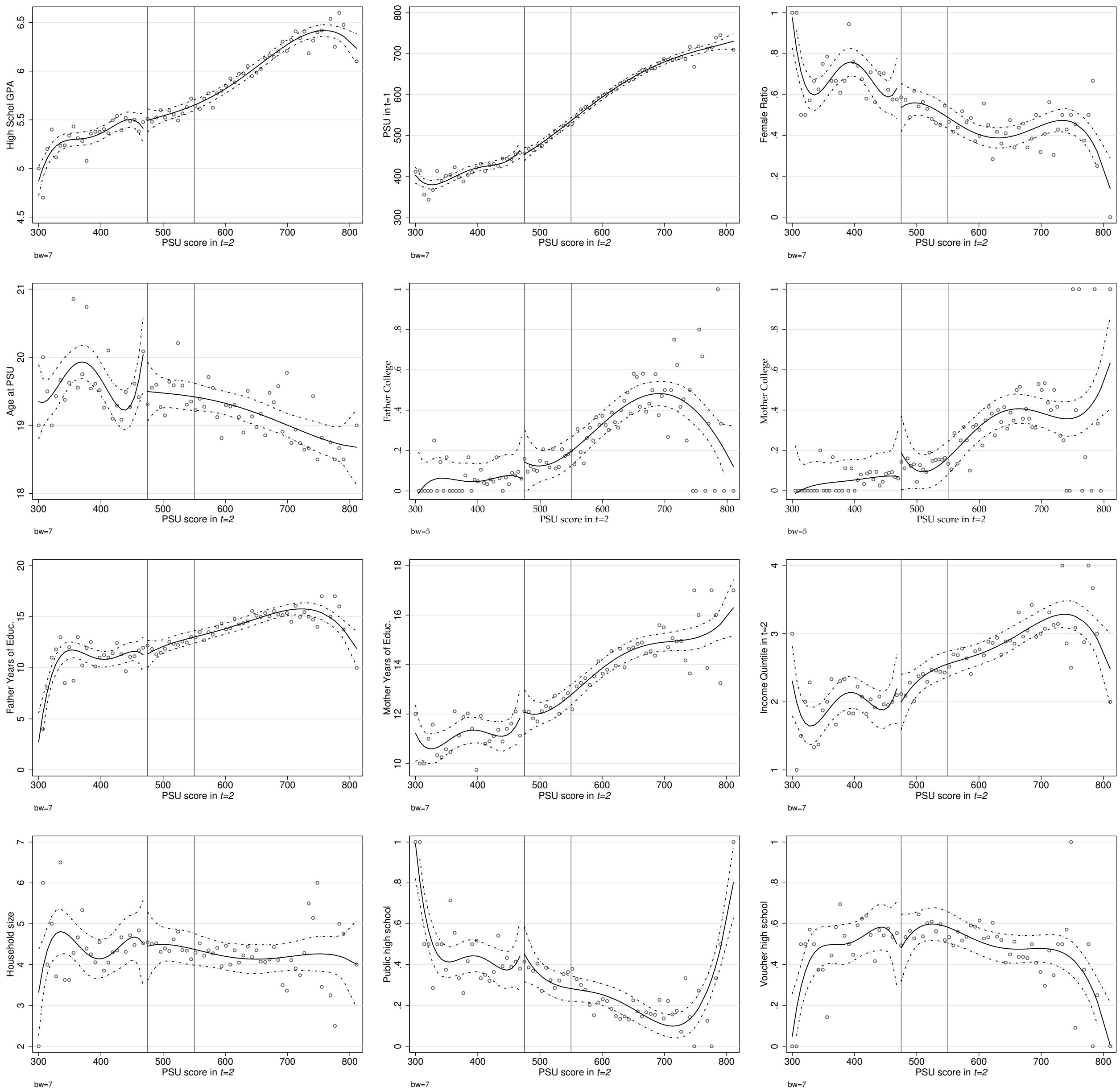
Note: Solid lines represent the estimate for each bandwidth used in the horizontal axis. Dotted lines represent 90 and 95% confidence intervals.

B.2 Balance of covariates

Figure B.2. Each dot represents the average of the characteristics, and the solid and dashed lines represent fitted values and the 95% confidence interval from a polynomial regression at each side of the threshold.

What we can conclude from the figures is that the imbalance in father college education (left middle figure) is due to the underlying noise of the data. The figure to the right, depicting mother college education, looks more worrying. However, the imbalance is likely to happen given the standard errors in the sample. The rest of the charts confirm that the problem is restricted to that group. More importantly, Table 2 and Figure B.2 indicate that these differences are not economically significant and, in general, are driven by a few outliers.

Figure B.2: Balance of Covariates. Graphical form in $t = 2$.



Note: Vertical lines at 475 and 550 correspond to the loan and Bicentenario scholarship cutoffs, respectively. Each bubble represents the average value of each variable for students in bins of 7 PSU-points. Solid lines correspond to a fitted fourth-order polynomial for the running variable at each side of the cutoff. Dotted lines represent 95% confidence intervals.

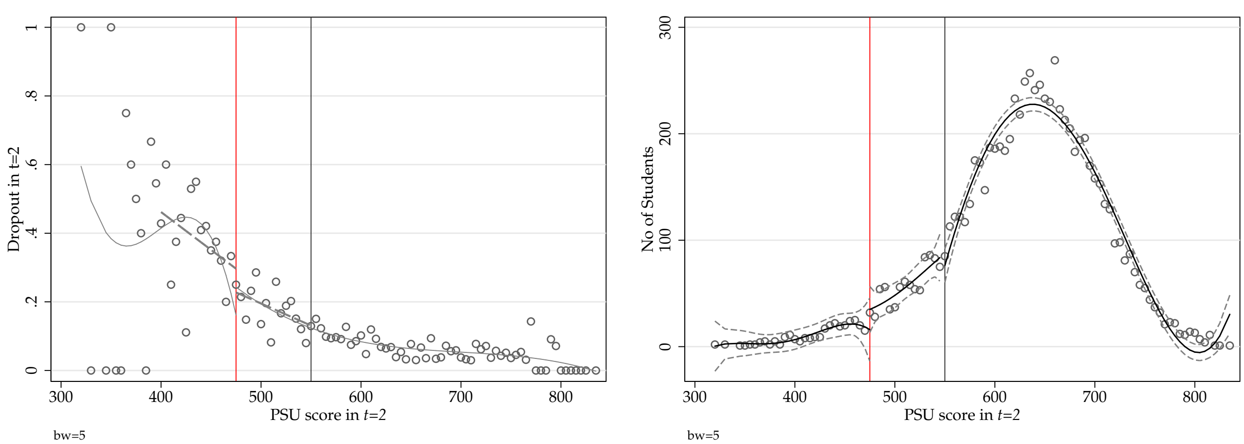
C Appendix. Placebo test

This appendix test other possible responses to crossing the eligibility cutoff. For example, given that financing is available for students above the cutoff, students may believe that the cutoff contains some information about who is prepared to succeed in college; or it could be that institutions are more likely to accept students above the cutoff because they believe that having funds to pay tuition would allow them to focus on their college work and increase the likelihood of graduation.

To test the existence of these behavioral responses, I use a group of similar students, i.e., ineligible students who enrolled in $t = 1$ who also retook the test, but who were not eligible for loans in $t = 2$– i.e., did not apply for loans in $t = 2$ or were classified in the richest income quintile. Given they are ineligible, loan access does not change at the cutoff, and therefore any change in the dropout rate would be a signal of responses not associated with credit constraints.

Figure C.1 shows the dropout rate in the second year in the left chart. The figure shows no apparent jump at the cutoff, indicating the absence of these type of responses.

Figure C.1: Placebo Test



Note: The figure at the top shows the dropout rate in the second year for students who did not apply for benefits (ineligible for loans) in either period. As in the main analysis, they also enrolled in college immediately after taking the PSU test. The figure at the bottom shows the empirical distribution of PSU scores for this sample.

The bottom figure shows the empirical distribution of PSU scores in $t = 2$ to show that the RD is valid for this group as well–i.e., it shows no manipulation of the running variable. This figure also shows that the number of students in this category is about a third of those used in Figure 2 (about 40 students at the cutoff), which may explain the extra noise. Secondly, it shows that the students across the cutoff– those who allow the identification –have lower ability than the mean of this group. This fact may be relevant if the decision to drop out depends on the relative position of one student with respect to her type.

Table C.1: Placebo Test. Dropout in $t = 2$ for ineligible

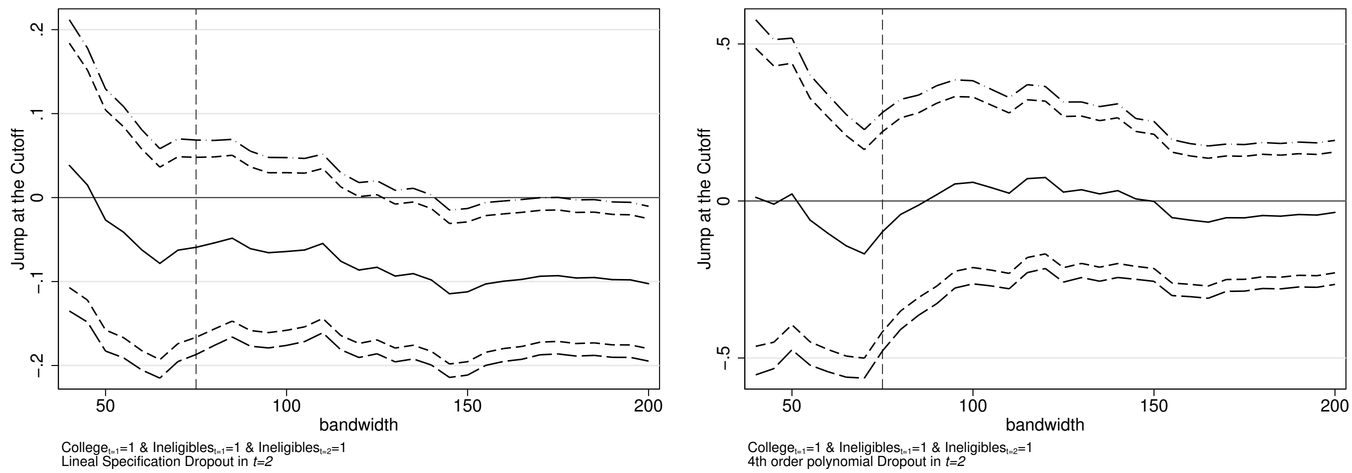
	P($E_{i2} = 0 \mid S_0, PSU_{i2}$)			P($V_{i2} = 1 \mid \cdot$)	P($M_{i2} = 1 \mid \cdot$)
	(1)	(2)	(3)	(4)	(5)
$1(PSU_2 \geqslant 475)$	-0.045	-0.060	-0.060	-0.057**	0.012
	(0.041)	(0.042)	(0.042)	(0.028)	(0.034)
Const.	0.29***	1.73	4.23	0.13***	0.16***
	(0.035)	(7.09)	(9.85)	(0.026)	(0.029)
Obs.	2,624	2,236	2,236	2,624	2,624
Controls		x	x		
Years FE			x		

Note: Estimation of equation 1 for dropout rates using the sample of ineligible–i.e., enrolled students who did not apply for loans in $t = 1$ retook the test in $t = 2$ but were not pre-selected for loans. All regressions use a local linear regression (f in equation 1) over a window of 60 PSU-points. Columns (4) to (5) decompose the effect in column (1) into the destination after dropping out. Robust standard error in parentheses. *: $p \leqslant 10\%$, **: $p \leqslant 5\%$, ***: $p \leqslant 1\%$

Table C.1 shows a formal analysis. The table follows the same structure as in Table 3. Column (1) shows that the dropout rate decreases eight percentage points at the cutoff, but that is not statistically different from zero. Columns (2) and (3) add covariates and years fixed effects with slight changes in the estimate but without modifying the conclusions of the first column. Columns (4) and (5) decompose this effects into the destination after dropping out, ruling out the possibility of canceling out effects.

Finally, Figure C.2 shows the sensitivity test for the choice of bandwidth. Both figures show the jump at the cutoff for independent regressions using bandwidths from two to 200 points. The left-figure use a local linear regression at each side and the right-figure uses a fourth-order polynomial at each side of the cutoff. In the left, we can observe that, for all reasonable bandwidths, the effect is not statistically different from zero. The effects turn to be significantly negative for bandwidths greater than 90 PSU-points. The right-figure shows that, for polynomial specifications, no bandwidth gives significant effects. These two figures suggest that there is no other type of responses. Nevertheless, this evidence is not as robust as expected.

Figure C.2: Sensitivity for the Placebo test. Jump for different bandwidths



Note: Solid lines represent the estimate for each bandwidth used in the horizontal axis. Dotted lines represent 90 and 95% confidence intervals.

D Appendix. Longer run effects

D.1 The Effects for the Fifth year

Table D.1 shows 2SLS estimates for dropout in the fifth year. Panel A decomposes the dropout rate into the chosen outside options, vocational education or the labor market, and Panel B shows the temporal decomposition. The non-enrollment rate in the fifth year is 58% for barely ineligible and decreases by 22 percentage points at the cutoff. Column (2) and (3) decompose the previous rate into the part of the non-enrollment gap in $t = 5$ that is attributable to people who stay enrolled in $t = 2$ but do not in $t = 5$, i.e., $p_{5,1,.,.,0}$; and those who dropped in $t = 2$ and did not enroll in $t = 5$, i.e., $p_{5,0,.,.,0}$. This decomposition highlights that no enrollment in second is the main driver of not enrollment in fifth, and the gap between eligible and ineligible. Among the ineligible, 41% dropped in second, and an extra 17% dropped in the following year. Among the eligibles, only 15% failed to enroll in the second, and 14% dropped in the next year. The option chosen after dropping out for these students was pretty similar.

Panel B confirms that the most significant period in the dropout process is the second year, which accounts for almost all the dropout observed to this point. 37% of the ineligible did not enroll in the second year and never went back to college, while only 15% among eligible students. In all other years, the rate of no enrollment was similar for both groups around 6% per year.

Table D.1: Dropout in the fifth year of college.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Decomposition by destination. 2SLS:								
	$p_{5,.,.,.,0}$	$p_{5,1,.,.,0}$	$p_{5,0,.,.,0}$	$p_{5,1,.,.,v}$	$p_{5,1,.,.,m}$	$p_{5,0,.,.,v}$	$p_{5,0,.,.,m}$	
Eligile for Loans	-0.19***	0.028	-0.22***	0.012	0.015	-0.15***	-0.071	
	(0.071)	(0.059)	(0.062)	(0.038)	(0.049)	(0.050)	(0.050)	
Const.	0.57***	0.19***	0.38***	0.068**	0.12***	0.23***	0.15***	
	(0.053)	(0.043)	(0.052)	(0.026)	(0.036)	(0.043)	(0.043)	
Obs.	869	869	869	869	869	869	869	
B. Temporal Decomposition. 2SLS:								
	$p_{5,0,0,0,0}$	$p_{5,0,0,1,0}$	$p_{5,0,1,0,0}$	$p_{5,1,0,0,0}$	$p_{5,0,1,1,0}$	$p_{5,1,0,1,0}$	$p_{5,1,1,0,0}$	$p_{5,1,1,1,0}$
Eligile for Loans	-0.19***	-0.0058	-0.0097	0.019	-0.015	-0.0046	0.0057	0.0077
	(0.061)	(0.0087)	(0.012)	(0.032)	(0.014)	(0.0033)	(0.033)	(0.044)
Const.	0.35***	0.0080	0.011	0.033	0.018	0.00068	0.060**	0.096***
	(0.051)	(0.0084)	(0.012)	(0.021)	(0.012)	(0.00051)	(0.025)	(0.032)
Obs.	869	869	869	869	869	869	869	869

Note: Estimation of equation 1 for dropout rates using the sample of analysis. First three columns use a local linear regression (f in equation 1) over a window of 60 PSU-points. Columns (4) to (6) uses a fourth-order polynomial f over the whole PSU support. Robust standard error in parentheses. *: $p \leqslant 10\%$, **: $p \leqslant 5\%$, ***: $p \leqslant 1\%$

D.2 Appendix: Longer-run Dropout. Reduced Forms

Table D.2: Dropout in $t = 4$. Reduced Form

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. Decomposition by destination. 2SLS:							
	$p_{4,.,.,0}$	$p_{4,1,.,0}$	$p_{4,0,.,0}$	$p_{4,1,.,v}$	$p_{4,1,.,m}$	$p_{4,0,.,v}$	$p_{4,0,.,m}$
$1(PSU_2 \geqslant 475)$	-0.11**	0.030	-0.14***	-0.011	0.041	-0.099**	-0.045
	(0.054)	(0.039)	(0.049)	(0.026)	(0.031)	(0.043)	(0.031)
Const.	0.48***	0.14***	0.34***	0.067***	0.075***	0.23***	0.11***
	(0.041)	(0.027)	(0.040)	(0.020)	(0.021)	(0.037)	(0.026)
Obs.	1,102	1,102	1,102	1,102	1,102	1,102	1,102
B. Temporal Decomposition. 2SLS:							
	$p_{4,.,.,0}$	$p_{4,0,0,0}$	$p_{4,0,1,0}$	$p_{4,1,0,0}$	$p_{4,1,1,0}$		
$1(PSU_2 \geqslant 475)$	-0.11**	-0.14***	0.000030	-0.0068	0.037		
	(0.054)	(0.048)	(0.011)	(0.027)	(0.030)		
Const.	0.48***	0.33***	0.0086	0.061***	0.081***		
	(0.041)	(0.040)	(0.0094)	(0.019)	(0.021)		
Obs.	1,102	1,102	1,102	1,102	1,102		
Controls		x	x				
Years FE			x				

Robust standard error in parentheses. $***: p \leqslant 1\%$; $**: p \leqslant 5\%$

Table D.3: Dropout in $t=5$. Reduced Form

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Decomposition by destination. 2SLS:								
	$p_{5,.,.,.,0}$	$p_{5,1,.,.,0}$	$p_{5,0,.,.,0}$	$p_{5,1,.,.,v}$	$p_{5,1,.,.,m}$	$p_{5,0,.,.,v}$	$p_{5,0,.,.,m}$	
$1(PSU_2 \geqslant 475)$	-0.16***	0.024	-0.19***	0.010	0.013	-0.13***	-0.061	
	(0.061)	(0.050)	(0.054)	(0.032)	(0.042)	(0.043)	(0.042)	
Const.	0.54***	0.19***	0.35***	0.070***	0.12***	0.21***	0.14***	
	(0.045)	(0.036)	(0.045)	(0.022)	(0.031)	(0.037)	(0.036)	
Obs.	869	869	869	869	869	869	869	
B. Temporal Decomposition. 2SLS:								
	$p_{5,0,0,0,0}$	$p_{5,0,0,1,0}$	$p_{5,0,1,0,0}$	$p_{5,1,0,0,0}$	$p_{5,0,1,1,0}$	$p_{5,1,0,1,0}$	$p_{5,1,1,0,0}$	$p_{5,1,1,1,0}$
$1(PSU_2 \geqslant 475)$	-0.16***	-0.0049	-0.0083	0.016	-0.013	-0.0039	0.0048	0.0066
	(0.053)	(0.0074)	(0.0099)	(0.027)	(0.012)	(0.0028)	(0.028)	(0.037)
Const.	0.32***	0.0071	0.0096	0.036**	0.015	1.4e-16	0.061***	0.097***
	(0.044)	(0.0071)	(0.0098)	(0.017)	(0.0098)	(.)	(0.021)	(0.027)
Obs.	869	869	869	869	869	869	869	869

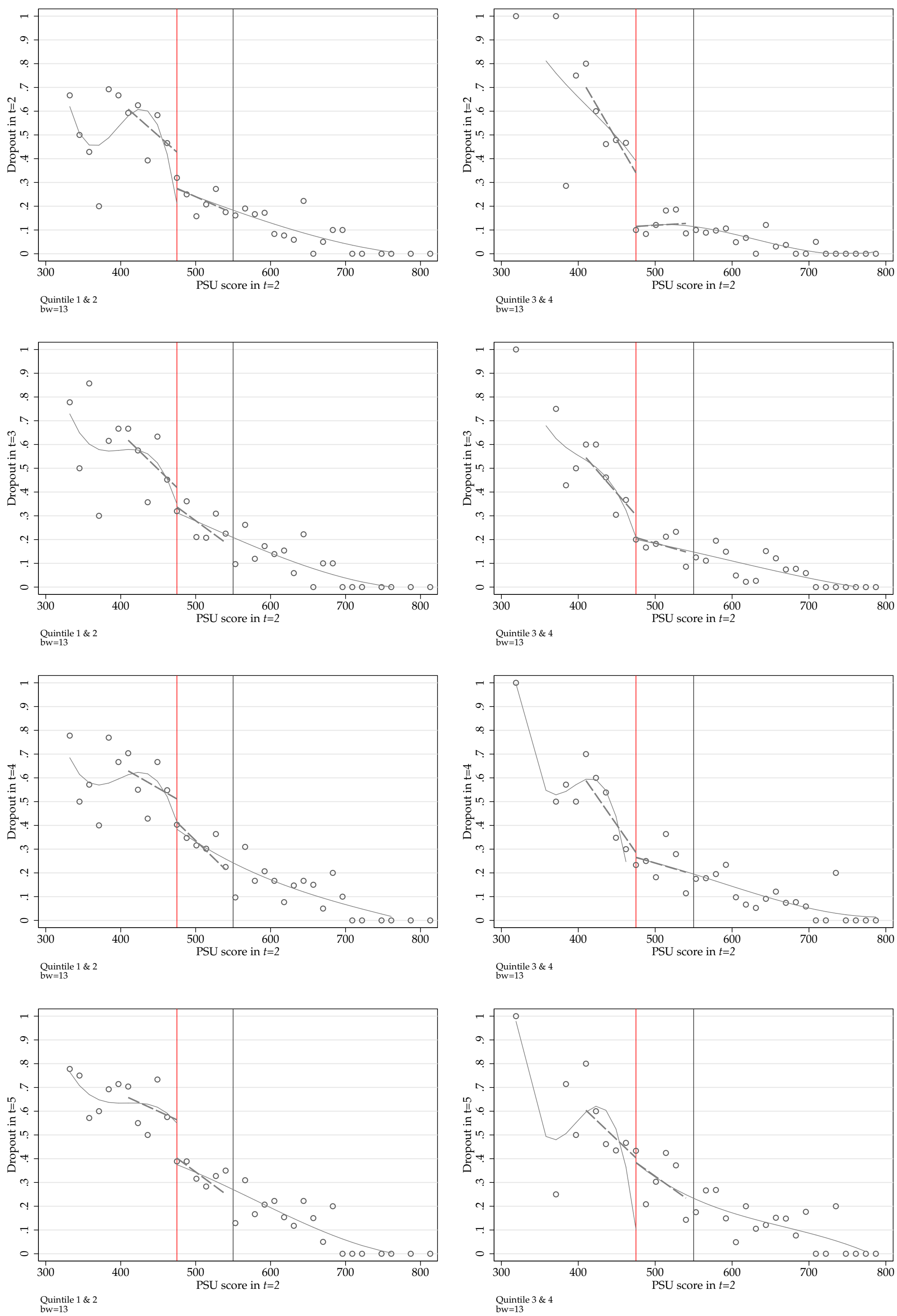
Robust standard error in parentheses. $***: p \leqslant 1\%$; $**: p \leqslant 5\%$

E Appendix: Testing other channels

Figure E.1 shows the analysis by income groups, using the classification made by the tax authority. Following the description of Figures 3, each dot represents averages for individuals within bins. Here, we use a width of 13-point bin to smooth the noise generated by smaller samples that consider fewer cohorts.

Figure E.1 shows the effects for different years of college. Each row corresponds to a different college year, starting with the second year on top until the fifth at the bottom row.

Figure E.1: Dropout in the second year of college by income quintile.



Note: Vertical lines at 475 and 550 correspond to the loan and scholarship cutoffs, respectively. Each bubble represents the average dropout rate among students in bins of 13 PSU-points. Solid lines represent a fitted fourth-order polynomial for the running variable at each side of the cutoff.

F Appendix: Testing other channels

Here I try to test the concern of the mechanism I described earlier: Students enroll for one year planning to default their bill if they do not get the loan in the next PSU attempt.

One way to test the importance of this mechanism is to see the dropout effects for students who are enrolled and are not eligible in $t = 1$ and $t = 2$, who retake the PSU test in $t = 3$. The first column shows that the effect is of similar size and strength despite the number of observations. Column (2) shows the proportion of students who were enrolled and not eligible in $t = 1$, but did not enroll in college in $t = 2$ (not enrolling means losing the right of using any benefit). The effect is a 50% reduction as before. The third column ads those who were not enrolled in any college (2-year or 4-year)

Table F.1: Different PSU processes PSU_3, PSU_4 and PSU_5

	$t = 3$			$t = 4$	$t = 5$
	(1)	(2)	(3)	(4)	(5)
$1(PSU_3 \geqslant 475)$	-0.14	-0.34***	-0.36***		
	(0.098)	(0.062)	(0.063)		
$1(PSU_4 \geqslant 475)$				-0.19**	
				(0.082)	
$1(PSU_5 \geqslant 475)$					-0.43***
					(0.12)
Const.	0.39***	0.64***	0.64***	0.56***	0.68***
	(0.078)	(0.046)	(0.047)	(0.060)	(0.089)
Obs.	342	763	736	462	199

Robust standard error in parentheses. $***: p \leqslant 1\%$; $**: p \leqslant 5\%$