A Loanly Road to Teaching:

Recruiting Talented Students Through Service Scholarships

Nicolás Rojas Souyet*

This version: November 9th, 2025

Abstract

Hiring skilled professionals for public service jobs in high-need areas is challenging, especially in teaching. A popular solution is service scholarships, college grants that convert into loans if recipients do not meet the agreed requirements to enroll, graduate, and work in targeted schools for a set number of years. Do these programs recruit additional talented students into teaching? I study the impact of Chile's free-tuition *Beca Vocación de Profesor* on attracting high-scoring students to teach in publicly funded schools. Using a fuzzy regression discontinuity design and 14 years of national data, I find that the program increased recipients' probability of teaching in such schools by 18 percentage points, a lasting effect beyond the service agreement. Drawing on a complier decomposition, I show that this impact was driven by increased enrollment, rather than by reduced dropout or improved transition into the teaching workforce. Consequently, 95% of recipients did not fulfill the agreement they signed, and the sizable grant converted into unmanageable debt. I discuss the challenge of addressing default—a persistent concern in loan-based aid programs that remains unaddressed in such programs.

*Columbia University. E-mail: nicolas.rojas@tc.columbia.edu. I thank Judith-Scott Clayton, Felipe González, Peter Bergman, Mary Laski, Dan Goldhaber, Sebastián Calónico, Tatiana Velasco, Alex Eble, Kirsten Mumma, Sandra Black, Randy Reback, Carolina Concha-Arriagada, Mariel Bedoya, Jack Willis, Zachary Bleemer, Jesse Bruhn, Sebastián Otero, Cristóbal Otero, Kiki Pop-Eleches, Jonah Rockoff, Viola Corradini, and the attendees of AEFP 2022, the NAEd 2024 Spring retreat, the Development Colloquium and Applied Economics at the Economics Department at Columbia and the Colloquium of Economics and Education at Teachers College for their useful comments and suggestions. An older working paper version title was: *Free education and the formation and sorting of talented teachers.*

1 Introduction

Many public service occupations struggle to fill vacancies in high-need sectors or regions. Among these occupations, education faces especially severe challenges: half of all U.S. schools experienced teacher shortages during the 2022–2023 school year (NCES, 2023), and the world faces an estimated deficit of 44 million primary and secondary teachers by 2030 (UNESCO, 2024). This problem is aggravated in low-income schools (OECD, 2020) and for more qualified teachers (UNESCO, 2024), a pressing issue given the substantial short- and long-term effects that more productive teachers have on students (Rivkin et al., 2005; Chetty et al., 2014; Araujo et al., 2016; Jackson, 2018). Recruiting and retaining teachers where they are needed most is difficult, but targeted incentives may help by correcting market failures that arise when compensation does not reflect their full societal value (Steele et al., 2010).

Service scholarships, grants aimed at attracting students to teach in high-need areas, are one popular solution. Recipients commit to enroll, graduate, and teach for a set number of years in targeted schools; if they fail to meet the terms of the agreement, the grant converts into a loan. The U.S. is the leading adopter of these policies, with at least 28 active service scholarship programs and new proposals to expand this model in response to the teacher shortage crisis[1]. Despite their popularity, evidence on the causal impact of these programs on recruitment into teaching—both during and after the service agreement—is scarce, and little is known about how generous but conditional grants shape recipients' paths (Steele et al., 2010). Do these programs recruit additional talented students into teaching in areas in need?

In this article, I study the impact of Chile's *Beca Vocación de Profesor* (BVP, or Teaching Vocational Scholarship)—a free-tuition service scholarship—on recruiting talented high-school graduates into teaching positions in publicly funded schools over time. Using 14 years of national administrative data, I estimate the program's effects with a fuzzy regression discontinuity design (RDD). To study the mechanisms that drive high-school graduates' choices into teaching in publicly funded schools, I compare the policy's impact on higher-education and teaching outcomes to counterfactual scenarios in which the policy only increases enrollment in teaching programs while reproducing low-persistence rates into teaching. In the presence of the threat of grant-to-loan conversion and the additional income provided by the scholarship[2], my estimates would be expected to differ from these counterfactuals by reflecting improved transitions into the teaching workforce.

Initiated in 2011, the BVP offers full tuition and fee subsidies to students enrolling in an eligible teaching major for the first time, provided they achieve an average score of 600 points or higher on Chile's college entrance exam, roughly the 85th percentile. In return, recipients must graduate and commit to teaching for three years in any publicly funded school within twelve years of graduation, or the generous grant converts

[1]The U.S. has 28 service scholarships at both the state and federal levels, and some of them have been running for more than 30 years. For a comparison with other countries and references on recent proposals, see Appendix A.1, Darling-Hammond et al. (2023) ; García et al. (2023) .

[2]or other program incentives, such a signaling a public service orientation.

into a sizable loan. The grant is up to 166% higher than what students in Chile can ask as a loan for teaching programs, a limit that is based on realistic labor-market returns for teaching program students (Eluchans, 2014; Beyer et al., 2015). The BVP has no income-contingent repayment option - a key feature of student loan programs in Chile, the U.S., and elsewhere (Lochner and Monge-Naranjo, 2016).

I address the selection bias problem arising from unobservable differences between recipients and non-recipients (such as intrinsic motivation for teaching or public service orientation) by exploiting the discontinuity generated by the 600 points eligibility threshold. Students close to this cutoff share similar observable and unobservable characteristics—except for their eligibility status—enabling identification with a fuzzy regression discontinuity design. I track the 2011 to 2013 cohorts from the national admissions test through their higher education enrollment, graduation, and into their teaching careers, with data concluding twelve years later. These cohorts were offered the same twelve-year agreement. Based on pre-policy data, the agreement seemed extremely difficult to fulfill, although this information was not disclosed to students.

I examine the mechanisms driving the program's impact on recruitment by constructing two counterfactual scenarios that help me assess how the agreement to enroll, graduate, and teach shapes the program's effects. First, I estimate what the policy's impact on higher-education and teaching outcomes would be if its only effect were to enroll additional students who behave like high-scoring teaching-program students who did not receive the BVP. Second, I estimate what the policy's impact would look like if all the recipients behaved as the type of recipient who would have studied teaching even in the absence of the policy, under the potential-outcome scenario in which they do not take the scholarship, following Abadie (2002). Both counterfactuals are based on estimates of the shares of students who react to the 600-point offer (the compliers) and are either additional teaching-program students (marginal students) or would have studied teaching regardless (inframarginal students). This complier decomposition takes advantage of the fact that recipients must study teaching—an outcome-monotonicity assumption embedded in the program's design—and builds upon the *supercompliers* framework proposed by Comey et al. (2024).

I divide my results into three parts: higher-education effects, teacher-market effects, and indebtedness by the end of the agreement. In my higher-education results, I find that the BVP increased the probability that students enroll in a teaching program—similar to Neilson et al. (2022)'s findings—and that they switch out of or drop out from a teaching program without graduating in another field. The probability that a recipient became a teaching-program student increased by 33 percentage points (p.p.), while the probability of switching out or dropping out without graduating in another field rose by 17 p.p. These results highlight how a service scholarship can significantly increase enrollment in teaching programs while exposing students to the risk of converting their grant into an unmanageable level of debt for three reasons. First, under the agreement the recipients signed[3], not finishing the program converts the sizable grant into a loan much larger than what students in Chile can normally borrow for teaching programs, given realistic labor-market returns.

[3] In 2017, the Ministry of Education implemented a drop-out insurance discussed throughout the paper.

Second, the BVP almost completely eliminated the probability that a recipient ever received a no-risk grant covering 80% of tuition, reducing it by 60 p.p. Third, the program did not increase the probability of ever enrolling in or dropping out of higher education, indicating that BVP incentives did not help students enter or remain in higher education.

My results align with both of my counterfactual scenarios, suggesting that the higher education effects primarily reflect the enrollment of additional students who exhibit similar low-persistence patterns to high-scoring untreated peers, rather than post-enrollment effects driven by the loan threat, income gains, or other program mechanisms. This additional enrollment comes from one-third of the compliers who are marginal in their decision to study teaching, while two-thirds would have pursued teaching regardless (inframarginals). Consequently, roughly three scholarships were required to recruit one additional teaching student near the 600-point cutoff. Interestingly, around 40 percent of the marginals chose teaching rather than waiting to re-take the test, consistent with present bias in educational choices (Lavecchia et al., 2016).

In my results on the teacher labor market, I find that the BVP policy succeeded in recruiting talented students who remained in publicly funded schools even after the agreement ended. It increased the probability of becoming a teacher at a publicly funded school for at least one year by about 18 p.p., around 14 p.p. for at least three years, and roughly 12 p.p. for contract hours beyond the service agreement period[4]. These dynamic effects are critical because productivity improves during the initial years of teaching (Rivkin et al., 2005; Papay and Kraft, 2015). Mirroring my higher-education effects, these results closely align with both of my counterfactual scenarios, suggesting once again that the program's effect primarily reflects the enrollment of additional students who exhibit similar low-persistence patterns to high-scoring untreated peers. High-scoring untreated peers who taught for at least one year in publicly funded schools generally remained in these positions. This explains why the agreement deadline does not induce even most of the recipients who teach in publicly funded schools to leave.

Finally, I study the impact of the policy on indebtedness and compare recipients' indebtedness arising from not fulfilling the program agreement to that arising from additional subsidized private or government loans, across different scenarios built to reflect the policy changes implemented after 2013. Although the introduction of drop-out insurance and gradual and partial repayment prevented nearly 95% of recipients from converting the entire grant into a loan and reduced the positive impact on indebtedness from 31 p.p. to a statistically insignificant –6.5 p.p., more than 55% of recipients still accumulated some form of debt by the end of the agreement. This outcome ultimately fell short of the "Studying for Free" promise in the original advertisements and aligns with the financial trajectories of high-scoring students in teaching programs before the program began.

This work contributes to three strands of the literature. First, it adds to the literature on the causal effects

[4]12 p.p considers changes in the agreement after 2013. I also find no significant or marginal small effects on teaching only one or two years in a publicly funded school or working at a private, unsubsidized school.

of recruitment and retention policies aimed at reducing teacher shortages in targeted areas by exploring an overlooked issue: the offer of sizable grants that can increase recruitment into teaching in targeted areas but may later convert into loans that exceed aid limits designed to prevent over-borrowing. This literature often finds that different strategies successfully recruit teachers in high-need areas including making disadvantaged schools more salient in teachers' recruitment platforms (Ajzenman et al., 2024), compensation increases for teachers working in rural schools (Bobba et al., 2021), implementing financial bonuses for high-performing teachers or teachers of targeted subjects in disadvantaged schools (Clotfelter et al., 2008; Elacqua et al., 2022), incentives for high-performing teachers to transfer to low-performing schools (Glazerman et al., 2013), loan forgiveness in "hard-to-staff" subjects such as STEM (Feng and Sass, 2018), and urban teaching residencies (Papay et al., 2012). My results show that, while service scholarships increase teacher recruitment in areas in need, recipients can be highly exposed to indebtedness that exceeds the safeguards built into student loan policies to prevent default. The theory on the optimal structure of loan arrangements in education indicates that borrowing limits and income-contingent repayment are essential for enhancing welfare in loan aid programs, as they address several market frictions (Lochner and Monge-Naranjo, 2016). Yet, service scholarships generally exclude such features, even though they are already part of student loan programs in places like Chile and the U.S. (Beyer et al., 2015).

Within the recruitment and retention literature, this paper contributes to the limited causal evidence on service scholarships that finds positive effects on teaching recruitment in areas in need (Steele et al., 2010; Neilson et al., 2022). I build upon Neilson et al. (2022), who also finds a positive impact of the BVP on teaching recruitment in any school, irrespective of its source of funding, using a sharp RDD at the same cutoff and eight years after enrollment. In contrast, I focus on recruitment into publicly funded schools, drawing on 14 years of data, including different types of schools. Because most students graduate in their sixth year or later, I can examine teaching outcomes across multiple years after graduation, including those after the agreement ends and across different school types. My fuzzy RDD and my counterfactuals provide further insights into who and how many students are recruited into the program and how teaching recruitment stems mainly from initial enrollment in teaching programs, reproducing sizable low-persistence patterns from college to teaching. These results are crucial to understanding how the agreement operates, an aspect of the BVP that is absent from Neilson et al. (2022).

Also within the service scholarship literature, this paper provides insights into how such programs shape higher education choices both within and beyond teaching programs. Existing research in Chile, the U.S., and China shows mixed results for increasing enrollment in teaching programs (Han and Xie, 2020; Peyton et al., 2022; Castro-Zarzur et al., 2022; Neilson et al., 2022), while showing no impact on course grades or college GPA (Zheng and Shi, 2024). I also build on Neilson et al. (2022), who finds positive effects of the BVP on immediate enrollment in teaching programs but no effect on higher education graduation. I extend this work by providing evidence on a more general choice set showing no evidence of short-term credit constraints

(Carneiro and Heckman, 2002), in contrast to other grant and loan policies in Chile that require much lower admission scores (Solis, 2017; Bucarey et al., 2020; Aguirre, 2021). This connects with my novel results on the substantial crowding out of no-risk grants by the BVP: high-scoring students already had access to higher education. By exploring several additional outcomes[5], such as graduation and on-time graduation in teaching programs, with my fuzzy RDD and my counterfactuals, I find no impact on trajectories beyond initial enrollment. Taken together with evidence from Chile, the U.S., and China (Han and Xie, 2020; Peyton et al., 2022; Castro-Zarzur et al., 2022; Neilson et al., 2022), my results suggest that service scholarships do not change educational trajectories beyond initial enrollment; their impact is limited to the decision to enter teaching programs.

Second, this study contributes to the ongoing debate on the merits of conditional loans versus conditional grants in financial aid. In a seminal paper, Field (2009) conducts an experiment at a law school showing that framing aid contracts of equivalent monetary value as a service scholarship rather than as a forgivable loan substantially increases enrollment and placement in the targeted jobs specified in the agreement, while reducing indebtedness. Caetano et al. (2019) and Evans et al. (2019) provide further experimental evidence showing that framing or labeling financial aid as a loan, rather than as a monetarily equivalent alternative, reduces enrollment—consistent with debt aversion. This paper complements this evidence by underscoring that even when conditional financial aid is framed and labeled as a service scholarship rather than a forgivable loan, it can attract a sizable number of students to an agreement that is extremely difficult to fulfill and can result in substantial, unmanageable debt.

Finally, I present a novel financial aid application along with two extensions of the *supercompliers* framework proposed by Comey et al. (2024), which address settings where treatment effects can reasonably be assumed to be null or unidirectional for specific binary outcomes. I build on this framework to distinguish and characterize marginal and inframarginal enrollment, offering valuable information for policy evaluation of teacher recruitment programs—an area of inquiry that dates back to the 1960s[6]—which can also be extended to other types of financial aid and to fields such as nursing, law, or medicine (Wiederspan, 2018; Dynarski et al., 2023). Policy evaluation has traditionally relied on hard-to-obtain and potentially biased surveys to estimate how many students would alter their choices in response to the program[7] (Rogers, 2009; Podolski and Kini, 2016; Valenzuela, 2020), whereas this application addresses that question using existing administrative data that may be readily available even while the program is unfolding. As extensions, I show how marginal estimates can be applied to assess program costs and simulate two types of counterfactuals that clarify the

[5]For example, I study on-time and overall graduation in teaching programs, overall higher-education graduation, enrollment and graduation in higher education at any year after the PSU, and outcomes such as dropping out of teaching programs, switching fields, and leaving higher education, among others. I also build upon Neilson et al. (2022) by considering the service component of the program and including financial aid outcomes, including the crowding out of grants and loans.

[6]"Federal staff worried that the 'goal of attracting primarily those who would not otherwise have gone into teaching has not so far been attempted'" Rogers (2009) quote of Lawrence and Austin 1967's Report on the Teach Corps between 1966 and 1967.

[7]This problem has been addressed using low–take-up surveys due to the difficulty of gathering information from past recipients. For example, see the Ministry of Finance survey for the BVP take-up is below 16 percent (Valenzuela, 2020). See Podolski and Kini (2016) for other reports summarized in the US. Low take-up warns about selection-bias; surveys asking for their beliefs of a sizable monetary benefit can lead to desirability bias.

mechanisms underlying long-term trajectories in fuzzy RDD coefficients. According to Manski (1997), there are numerous applied settings in economics where researchers can confidently assert outcome monotonicity (e.g., a subsidy affecting service demand), enabling the extraction of valuable information, as demonstrated in this case[8].

The rest of the article is organized as follows. Section 2 describes the policy context and the data. Section 3 details my empirical strategy. Section 4 presents my main results. Section 5 concludes and discusses how the result from this paper highlights three key challenges in service scholarship design: the crowding out of no-risk grants, a loan component design that addresses over-borrowing concerns, and the need to provide information to potential recipients.

2 Policy Context and Data

This section outlines the program, highlighting the tension between attracting students with a sizable grant and establishing a feasible work agreement under the policy. It then describes the data sources used and provides a summary of the student sample studied in this project.

2.1 Policy Context

As in many other countries, Chile has struggled to attract the most talented students into the teaching profession (OECD, 2018). In the early 2000s, most students in teaching programs came from the lower half of the distribution in higher education admissions. Additionally, the curricula of teaching programs were difficult to modify, and the least qualified teachers were more likely to work in the most vulnerable schools. In response to this situation, the Ministry of Education implemented a service scholarship called *Beca Vocación de Profesor* (BVP) (or Teacher Vocation Scholarship) in 2011, aiming to attract the most skilled high school students into teaching (Beyer et al., 2010; Claro et al., 2013).

The program targeted high-scoring students. To be eligible, students had to be accepted into a teaching program for the first time and achieve an average score of at least 600 points in math and language on the PSU—Chile's centralized higher education admissions test—approximately the 85th percentile. High school graduates with over 580 points could still receive the scholarship if they graduated from a public school and ranked in the top 5% of their GPA cohort—a strict requirement[9]. Neilson et al. (2022) shows a robust,

[8] Labor economics reached a consensus that the Social Security Disability program in the US (SSDI) is a major force behind a decrease in employment, and the debate is on the magnitude of the impact rather than its direction (Abraham and Kearney, 2020; Maestas et al., 2013). In environmental economics, the debate on negative externalities' impact on health outcomes focuses on the size and timing of the effect rather than a positive or negative effect (Ebenstein et al., 2016; Currie et al., 2014; Chay and Greenstone, 2003). In heath economics, the effect of following breast cancer screening recommendations is expected to be null or positive on the probability of finding an in-situ tumor and reducing the average tumor size of a cancer diagnostic Einav et al. (2020) . Finally, in some information assistance programs in education, it can be difficult to argue that information could negatively affect some outcomes. For example, the study of the effect of smart-matching platforms to warn parents about misplacement risk on its impact on nonplacement (Arteaga et al., 2022) or the effect of student support for financial assistance on college participation Bettinger et al. (2012).

[9] Additional benefits were available above the 700 and 720 threshold, reaching only a few students, having no effect on

positive, and concave correlation between teachers' PSU scores and several short- and long-run teaching productivity outcomes, considering teachers' college value-added, which supports the teaching potential of high-scoring high school graduates.

On its face, the offer was generous. It covered full tuition and enrollment fees, where the average yearly tuition cost for teaching majors was approximately 3,200 USD[10] per year[11], about 15% of a student's family income for those around 600 points[12]. The funds were available only for the official duration of on-time graduation, which averages 4.8 years, and yearly renewal required passing at least 70% of the courses as an incentive for timely graduation.

In return, students signed a promissory note agreeing to repay the scholarship through a teaching service. If they failed to fulfill this obligation, the generous grant would convert into a sizable loan. The grant is on average 24% and up to 166% higher than what students in Chile can ask as a loan for teaching programs[13], a limit that is based on realistic labor-market returns for teaching program students (Eluchans, 2014; Beyer et al., 2015). The loan had to be repaid up front, adjusted for inflation, with no interest. The agreement required students to graduate and teach for three years, for at least 30 hours a week in any publicly funded school[14] (accounting for 92% of total enrollment), within 12 years after signing the agreement. Given that teaching programs lasted an average of 4.8 years, the loan could cost up to 15,000 USD—about 75% of a student's yearly family income[15] for those with 600 points. This is comparable to conditional aid programs in the US (e.g., up to 4,000 USD per year for the TEACH grant), but equivalent to 45.7 minimum wages in Chile, a much lower-income country.

The teaching context in Chile conflicted with the agreement in three critical ways: high dropout rates from teaching programs, graduates not teaching in publicly funded schools, and the strictness of the teaching requirement. In Figure 1, I depict the teacher pipeline for high-scoring students in the year before the policy was implemented. Importantly, this cohort could not have anticipated the policy, as it was discussed after they enrolled. Notably, 33% of students in teaching programs did not graduate, and close to 50% never taught in a publicly funded school within twelve years after enrollment. For this group, fulfilling even one year of the teaching requirement would have been as risky as tossing a coin. Remarkably, the students who

enrollment (Neilson et al., 2022; Ramaciotti et al., 2020). This will not be covered in this work.

[10]Costs are lower than other majors for students having more than 600 points, but similar to teacher majors in non-eligible institutions for the same group. For more information, see Table B6 in the Online Appendix B.2.

[11]All amounts are expressed in December 2015 dollars, using an exchange rate of 707 pesos per dollar.

[12]I estimate that students with 600 points are above the 75th percentile of income among students who attended higher education using HE using PSU and MINEDUC Data. Then, I estimate the 75th percentile of income families with at least one child in higher education using the CASEN household survey of 2015.

[13]I calculate these estimates based on MINEDUC data on the loan limits called "arancel de referencia". Both results come from the distribution of the loan limits that is weighted by the program enrollment of BVP recipients in my sample. 24% is the ratio of tuition cost, which was fully covered by the BVP, minus the loan aid cap, divided by the tuition cost. Note that according to a MINEDUC report (Eluchans, 2014), the loan limits should be even lower for teaching programs.

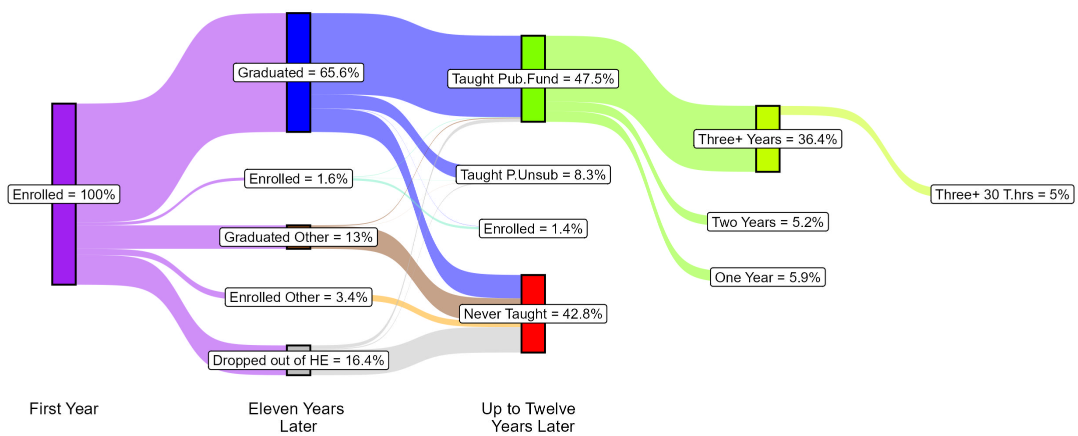
[14]This includes K-12, Pre-K, adult and special education institutions; however almost all of recipients studied and taught in K-12 (Ramaciotti et al., 2020).

[15]In December 2015, dollars and the exchange rate of 707 Chilean pesos per dollar. 15,000 USD amounts to 10,605,000 CLP. I estimate that students with 600 points were above the 75th percentile of income among students attending higher education with MINEDUC and PSU Data. Then, I estimate the 75th percentile of income families with at least one child in higher education using the CASEN, the national household survey of 2015.

taught at publicly funded schools continued teaching in these positions. However, what is most striking is that only 5% of students could have met the original teaching requirement.

Figure 1 does not describe a circumstantial trajectory. In Figure B1 in the Online Appendix B.1, I use the first PSU cohort seven years before the policy was implemented, finding a substantially similar pattern. No information about these risks was provided to the students.

Figure 1: High-Scoring Students' Teaching Program Trajectory
One Year Before the Policy Implementation



Notes. This figure illustrates that the original agreement was extremely hard to fulfill for high-scoring teaching program students the year before the implementation of the program. In all columns, outcomes are prioritized from top to bottom. For example, if someone graduated from a teaching program, this will count as "Graduated" regardless of whether the student also enrolled in or graduated from another teaching program or a non-teaching program. High-Scoring Students: Students scoring $\geq$ 600 at admissions, HE: Higher Education, Other: Programs $\neq$ from Teaching, T.hrs: Teaching hours per year P.Unsub: Private unsubsidized schools

The incompatibility of the scholarship with other financial aid sources made the program less generous than it initially appeared. Since recipients have their full tuition covered, they could not apply for a no-risk grant while receiving the scholarship or a subsidized loan for repayment if the grant converted into a loan. The most important grants were the Bicentenario (BC) and the Juan Gomez Millas (JGM) scholarships. The BC scholarship covered, on average, around 81% of the yearly tuition that the BVP covered, while the JGM covered 1,600 USD annually. The two college loan programs offered by the Ministry of Education are the Traditional University Loan (TUL) and the State Guaranteed Loan (SGL), which only cover tuition fees. All these financial aid programs had significantly lower eligibility cutoffs than the BVP: 475 points for the loans and 550 points for the grants. Their coverage was substantial: 50% of all PSU test-takers used one of these programs in their first year of enrollment between 2011 and 2015, with around 25% receiving at least one grant and 25% taking at least a loan. As a result, the full BVP grant could convert into a loan much larger than the limits set by the loan aid's policy to prevent over-borrowing -capped at less than 80% of tuition- under Chilean loans regulations (Eluchans, 2014; Beyer et al., 2015).

The original agreement was short-lived. Between October 2013 and March 2023, the decree was amended ten times, gradually relaxing its terms through five main changes. One key amendment introduced a gradual repayment scheme, supplementing the strict rule of teaching 30 hours per week for three years with an ambiguous 'equivalent trajectory' alternative (October 2013). This ambiguity was later resolved in January 2020 by clarifying that the service could be fulfilled by working the equivalent hours of three years in a 44-hour-per-week full-time contract. Two other significant changes included the introduction of partial repayment in January 2019, which allowed recipients to repay only the unmet portion of the teaching obligations, and the introduction of full drop-out insurance in January 2017, ensuring that recipients who did not graduate would not need to repay anything. Furthermore, the decree extended the original 12-year window for fulfilling the service requirement to a more flexible seven-year period after graduation (June 2017). Finally, the service repayment for priority geographic areas and specific subjects -such as religion, visual arts, music- was reduced to less than three years (October 2013, September 2015, June 2017, January 2020, and June 2021).

Despite this environment, the BVP was advertised as a free tuition scholarship for students over 600 points (see official advertisement in Figure A1 in Appendix A.2). This is consistent with how the Chilean Ministry of Education website describes a grant or a scholarship: "Unlike a loan, you do not need to repay these funds upon graduation or obtaining your degree." (MINEDUC, 2017).

Finally, the program established a relevant institutional constraint. To increase the number of high-scoring teaching majors, institutions were eligible for the BVP only if less than 15% of their teaching majors programs had scored below 500 points on the PSU, and if they were accredited by the government for at least two years. In Table B7 of the Online Appendix B.2 , I show that this restriction may have substantially increased the scores of teaching program students while strongly segregating the demand by admission scores. However, even before the program started, almost no student scoring 600 points or higher enrolled in an ineligible teaching program (see Figure A2 in Appendix A.2). In this paper, I focus only on the effect of the scholarship on students near the 600-point threshold, given that the scholarship program is active. A more general equilibrium effect is outside the scope of this analysis.

2.2 Data

I use six data sources to analyze the effects of the BVP. With these data, I follow the high school graduates who took the PSU, Chile's nationwide centralized admissions test, during the first three years of the program (2011 to 2013). Within these cohorts, the scholarship recipients signed the same agreement before any policy changes.

The first source of data is individual admissions data provided by the DEMRE, the agency in charge of college admissions. This includes PSU test scores, as well as socioeconomic and demographic characteristics collected in a survey administered prior to the test day. Two key elements of this data are the math and language scores. The average of both scores determines whether a student is eligible for the program, along with being admitted in a teaching program. Students apply to a major after receiving their scores.

The second source of data is MINEDUC's financial aid assignment data from 2011 to 2024. These datasets include individual yearly take-up on the BVP, as well as other government grants and loans, including the BC, the JGM, other grants, and the TUL loan program. For the BVP grant data, I only consider recruitment in the admission year corresponding to the test. PSU scores can also be used the following year, but not in subsequent years. The 600-point threshold induced a small share of students to take the scholarship in the year after enrollment (0.4% of all test-takers). The treatment variable does not include these cases and is defined as taking the scholarship in the same year as the test-taking process. This approach ensures that the entire BVP sample signed the same agreement before any policy changes, allows for the use of data up to 12 years after graduation, and facilitates building two counterfactuals to evaluate fuzzy results.

The third source of data is MINEDUC's individual national higher education files from 2011 to 2024. These datasets include yearly enrollment and graduation data by major, enabling the measurement of switching out of education majors and dropping out of higher education within 12 years of enrolling in the service scholarship program.

The fourth data source is MINEDUC's *Cargos Docentes* individual educational staff registry from 2011 to 2024. This dataset includes annual records of teaching contract hours and instructional hours by school type. I use this data to measure teaching trajectories within and outside of the agreement over 12 years following the receipt of the scholarship. Specifically, I construct teaching measures as teaching-status indicators, interacting them with enrollment in a teaching program during the admission year of test-taking. This analysis focuses solely on students who both received the scholarship and enrolled in a teaching program in the same year they took the test[16].

The fifth source of data is yearly individual take-up data from state guaranteed loans (SGL) provided by private banks, the SGL program. This data is provided by the INGRESA commission. I have records spanning from 2011 to 2022, covering ten years after the enrollment of my latest PSU cohort. Using data from INGRESA and MINEDUC financial aid programs, I study the crowding out of other sources of financial aid and the level of indebtedness at the end of the program, under the original agreement.

The sixth source of data is individual repayment records from the BVP program through April 2024. I use this data in the last part of the results in this article to illustrate repayment in relation to indebtedness.

2.3 Sample Characteristics

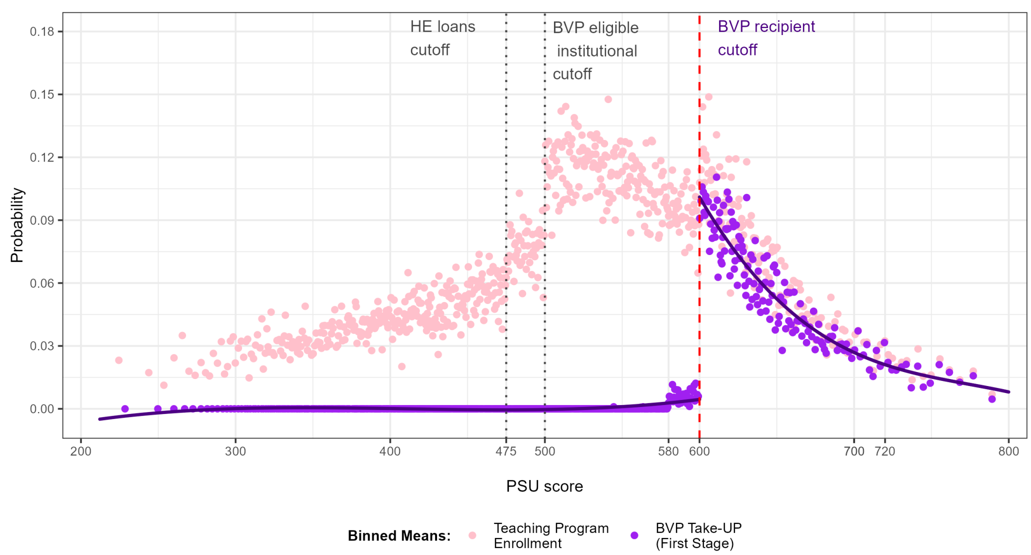
My empirical strategy relies on the program's admission cutoff at 600 points. Crossing this threshold automatically granted students access to the scholarship if they were accepted and enrolled in an eligible teaching program for the first time. Students were aware of their scores during the higher education program

[16]The results do not change substantially when interactions are excluded. Including them complicates the interpretation of results, especially for agreements signed after the policy changes in late 2013. Results without interactions and those using multiple years are briefly considered, regardless of higher-education outcomes, by including 'ever enrolled' or 'ever graduated' indicators to provide some perspective on this issue.

application process, which increased the probability of applying to and enrolling in the BVP program in the test academic year, but not later, in most cases. Importantly, even above the 600-point threshold, students tended to come from more disadvantaged backgrounds.

The program widely recruited high-scoring students, particularly those near the 600-point threshold. Using BVP and higher education enrollment data along with DEMRE's admission data, Figure 2 contrasts policy take-up with teaching program enrollment at different PSU score levels, leading to three key takeaways. First, there is a clear, sizable jump in take-up at 600 points, with few students taking the program between 580 and 600 points[17]. Close to 10% of all the test-takers in Chile with scores slightly above 600 points took the scholarship. Second, there is a significant overlap between policy take-up and teaching program enrollment for scores above 600 points. Specifically, 84% of teaching program students with more than 600 points enrolled in the BVP. Third, teaching program enrollment sharply declined above 500 points, with the area around 600 points showing the highest concentration of students taking the scholarship -a relevant fact for external validity. On average, 5.7% of test-takers with scores above 600 points enrolled in a teaching program, which is lower than the nearly 10% observed at 600 points.

Figure 2: Policy Take-Up (First Stage) vs Teaching Program Enrollment



Notes. (1) This figure illustrates a clear and large jump in take-up at 600 points (first-stage), the substantial overlap between enrollment and take-up, and that the highest take-up density is at 600 points. To see the jump in Teaching Program Enrollment, look at RDD plots in the Online Appendix, Section B. Since 84% of the teaching program enrollment are recipients, but not 100%, take-up mean-binned are slightly below, covering the discontinuity of enrolling in teaching programs. (2) N = 731,269 ; Mass Points = 1,263 (3) Data: DEMRE PSU test-takers from 2011 to 2013 merged with MINEDUC Higher Education Data. (4) Non-parametric means: I use an IMSE-optimal quantile-spaced method to estimate non-parametric means for different PSU scores using polynomial regression implemented in the rdplot R command described by Calonico et al. (2015a) and Calonico et al. (2015b).

[17] To see the jump in Teaching Program Enrollment, look at RDD plots in the Online Appendix Section B. Since 84% of the teaching program enrollment are recipients, but not 100%, take-up mean-binned are slightly below, covering the discontinuity of enrolling at teaching programs.

Table 1: BVP Offer Threshold Impact on BVP Take-up Over Time

	Immediate Take-up	Further Take-up			
	Same Academic Year	T+1 Academic Year	T+2 Academic Year	T+3 Academic Year	T+4 Academic Year
Take-up Effect	.092***	.002**	0	0	0
	(.002)	(.001)	(0.000)	(0.000)	(0.000)

Notes. (1) This table uses a sharp RDD to estimate the effect of crossing the 600-point BVP threshold for the admission year on joining the program in the admission year or in subsequent years. The main result is that the BVP offer almost entirely affects take-up in the year the test was taken (immediate take-up), but not later (further take-up). (2) Data: DEMRE PSU test-takers from 2011 to 2013 merged with MINEDUC Higher Education Data. (3) All the results are Linear Local Regressions (LLR) mean estimates at the 600 points cutoff with a uniform kernel. Heteroscedasticity robust standard errors in parentheses. *p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01 of estimates considering non-parametric bias and heteroscedasticity using the robust estimates of Calonico et al. (2014).

Table 2: Sample Mean Characteristics

	Full Sample Mean		LLR Mean at 600 points				
						Decomposition	
	All Obs N=731,269 $\overline{PSU}$=503 (1)	All BVP N=7,742 $\overline{PSU}$=635 (2)	All Stdnts at 600 points (3)	Always TP.Stdnts (4)	All Compliers (5)	Marginal TP.Stdnts Complier (6)	Inf.marginal TP.Stdnts Complier (7)
Share of Students at 600 points			100%	8.9%	9.2%^{c}	3.1%	6%
				[.083;.095]	[.087;.096]	[.023;.039]	[.053;.067]
Female	.527	.544	.494	.588	.557	.448	.614
	[.526;.528]	[.533;.555]	[.491;.499]	[.566;.606]	[.535;.59]	[.274;.59]	[.573;.716]
Age	18.91	19.73	18.75	19.29	19.86	21.12	19.2
	[18.9;18.92]	[19.65;19.82]	[18.75;18.78]	[19.16;19.44]	[19.69;20.15]	[19.85;22.42]	[18.8;19.72]
Private Health Insurance	.215	.292	.358	.252	.248	.216	.264
	[.214;.216]	[.282;.302]	[.352;.361]	[.237;.276]	[.217;.265]	[.086;.302]	[.221;.299]
Lower Income ($\approx <$Q50)	.515	.389	.338	.426	.436	.516	.394
	[.513;.516]	[.378;.4]	[.332;.34]	[.399;.446]	[.414;.471]	[.398;.672]	[.324;.446]
Mother Education	11.47	12.63	12.92	12.25	12.24	12.35	12.18
	[11.47;11.48]	[12.55;12.7]	[12.91;12.96]	[12.11;12.4]	[11.98;12.38]	[11.41;13.25]	[11.7;12.44]
Public High School	.328	.289	.226	.273	.275	.283	.27
	[.327;.329]	[.279;.299]	[.219;.226]	[.253;.293]	[.247;.293]	[.154;.376]	[.237;.322]
Voucher Private School	.564	.549	.577	.621	.615	.596	.62
	[.563;.565]	[.538;.56]	[.577;.589]	[.6;.637]	[.597;.654]	[.495;.75]	[.57;.68]
Private Unsubsidized School	.108	.162	.192	.111	.109	.133	.1
	[.107;.109]	[.154;.171]	[.184;.194]	[.099;.127]	[.087;.123]	[.042;.198]	[.07;.12]
N° of PSU Test Attempts	1.31	1.62	1.5	1.63	1.63	1.75	1.57
	[1.31;1.31]	[1.61;1.64]	[1.48;1.5]	[1.6;1.66]	[1.58;1.67]	[1.51;1.96]	[1.47;1.67]

Notes. (a): This table shows pre-treatment mean characteristics using DEMRE PSU pre-test survey and DEMRE PSU data-sets. The main goal is to describe all the PSU test-takers and BVP recipients, along with the characteristics of the Compliers at 600, the unit of analysis in this paper. The Appendix includes an expanded table with the always takers and more characteristics. **(b):** Columns (3) to (7) are Linear Local Regressions (LLR) mean estimates at the 600 points cutoff based on Calonico et al. (2014) LLR method. (2): LLR non-parametric mean based on Calonico et al. (2018) (3): Right-side of 600 cutoff of test-takers not enrolled at a Teaching Program (TP). (4): Left side of the cutoff of test-takers enrolled at a TP. (5): BVP recipients to the left of the cutoff. (6) BVP Compliers mean characteristic. (7): BVP Compliers induced to enroll at a TP by the program at a 600 mean characteristic. (8): BVP Compliers induced to take the program but who would enroll at a TP regardless of the program's mean characteristic. (7) and (8) are based on adapting Comey et al. (2024) to a Fuzzy RDD setting. **(c):** The share of marginal compliers equals 3.148695% and compliers 9.154834% **(d):** Income variables come from brackets $\{1,2\}$, $\{3,4\}$, $\{5,\ldots,12\}$ on the DEMRE family income question. **(e):** [.,.]: are 95% confidence intervals (CI). In (1) Standard normal CI. In (2) non-parametric robust CI using Calonico et al. (2018) . In (3) - (6) non-parametric robust CI using Calonico et al. (2014) . In (7) robust CI using a wild-bootstrap procedure adapted from He and Bartalotti (2020).

Crossing the program's admission cutoff at 600 points increased enrollment in the BVP program during the test academic year but had little to no effect in subsequent years. This supports the choice of following 2011 to 2013 take-up to identify the recruitment effects on students who signed the first version of the service agreement. The 600-point threshold did not induce recruitment in subsequent versions of the program. Table 1 shows the impact of crossing the admissions threshold at 600 points on BVP take-up during the test academic year and in later years. The first column estimates the first-stage jump in Figure 2 . Beyond the first year, there is almost no effect, except for a small but significant and precise impact in the second year. This result is consistent with the PSU's validity for students using the test either in the same academic year or the following year, but not in subsequent years.

As expected, students with 600 points came from more advantaged backgrounds than the average test-taker, scoring 0.9 standard deviations (SDs) above the average admission score of 507 points. Table 2 describes test-takers using DEMRE's pre-test survey. Columns (1) and (2) provide sample mean characteristics of all test-takers and all BVP recipients, along with 95% confidence intervals (CIs). The remaining columns present local-linear means at 600 points, along with bias-corrected non-parametric CIs. An average student scoring 600 points in column (3), clearly came from more advantaged backgrounds compared to the average test-taker with a score of 507 points in column (1), across various metrics. For example, the probability of a student scoring 600 points coming from a lower-income family (33.8 p.p) was 34% lower than that for the average test-taker (51.5 p.p), and their probability of having health insurance (35.8 p.p) was almost 70% higher. Students at the 600-point level were also twice as likely to come from private unsubsidized schools (19.2 p.p) compared to the average test-taker[18].

However, the average BVP recipient came from a more disadvantaged background than the average student scoring 600 points, even though their average score was 37 points higher than 600, equivalent to 0.34 SDs. For example, scholarship recipients were 5 p.p more likely to come from a lower-income family (15% of the baseline) and 6.6 p.p less likely to have private health insurance (18% of the baseline).

These comparisons are important for understanding recipients' risk of failing to comply with the agreement and for refining criticisms of using admission scores in this policy, given the strong correlation between socioeconomic status and admission scores. (Ramaciotti et al., 2020).

3 Empirical Strategy

I use the 600-point admission cutoff to identify the causal effect of the scholarship on higher education and teaching market outcomes among test-takers. Leveraging the fact that students near the 600-point eligibility threshold share similar observable and unobservable characteristics—except for their eligibility status- I implement a fuzzy regression discontinuity design (RDD). This strategy helps me overcome the selection bias problem arising from unobservable differences between recipients and non-recipients.

I also provide further insights to assess the practical (or economic) significance of the fuzzy estimates (Lipsey et al., 2012). I take advantage of the program's requirement that recipients must study teaching to decompose the share of recipients induced to take up the scholarship at the 600-point threshold — referred to as compliers — into those who are marginal students in teaching programs and those who would have

[18] Table A4 and Table A4 in the Appendix present an expanded table with more characteristics and confidence intervals for the difference between different subgroups.

studied teaching regardless, but without the scholarship. Using this decomposition, I characterize the two groups, compare them to the rest of the sample, and establish two counterfactuals along with the treatment effect baseline to assess whether the fuzzy estimates after teaching program enrollment could come only from enrolling more students similar to high-scoring untreated teaching program students, rather than from the loan threat or the income increase provided by the program.

3.1 Method: Fuzzy RDD

It is not possible to identify the average treatment effect of the program because of selection bias. Recipients may have higher motivation, more optimistic beliefs about teaching, or other relevant unobservable characteristics compared to non-recipients, making it difficult to disentangle enrollment in the program from unmeasured differences between students. However, near the 600-point cutoff, scoring slightly above or below the threshold can be considered random and not due to unobservable characteristics. This enables the estimation of the average treatment effect for the complier group at 600 points. The following equation describes the main estimand of interest for this work:

$$\beta_{FRD} \equiv E[Y(1,1) - Y(0,0)|S = 600, C] \tag{1}$$

β_{FRD} represents the average change in outcome Y for a complier student scoring 600 points who is treated and induced to take the treatment, compared to a complier student who is not treated and not induced to enroll in the program. Using potential outcomes notation, $Y(Z,D)$ represents the potential outcome under score eligibility Z and program take-up D. Take-up is implemented as immediate take-up -taking-up the program right after admissions- for outcomes that follow immediate enrollment, and ever take-up -taking-up the program right after admissions- for outcomes that do not follow immediate enrollment (eg. enrolling in higher education in any year after the PSU). $Z = 0$ indicates a score below 600 points, and $Z = 1$ indicates a score equal to or above 600 points. $D = 0$ indicates no scholarship take-up, and $D = 1$ indicates scholarship take-up. Compliers (C) are students who take up the scholarship only if they score above 600 points. S is the PSU score.

Under four assumptions, β_{FRD} can be identified as follows:

$$\beta_{FRD} = \frac{\lim_{s \to 600^+} E[Y \mid S = s] - \lim_{s \to 600^-} E[Y \mid S = s]}{\lim_{s \to 600^+} E[D \mid S = s] - \lim_{s \to 600^-} E[D \mid S = s]} \tag{2}$$

The first assumption is the continuity of the potential outcome mean functions, $E[Y(1,1) \mid S = s]$ and $E[Y(0,0) \mid S = s]$, and the potential take-up mean functions, $E[D(1) \mid S = s]$ and $E[D(0) \mid S = s]$, at $S = 600$. This implies that the observable average take-up or outcome of a student with slightly less than 600 points would be equal to the observable average take-up or outcome of a student with slightly more than 600 points, had they been assigned a score slightly lower than 600 points. This assumption makes both groups around 600 points comparable except for their treatment status, overcoming the selection-bias problem around the program cutoff. It also serves as an independence assumption, in which the average potential outcome and the average potential take-up do not depend on whether a student is slightly above or below 600 points.

The second assumption is the continuity of the potential outcome mean functions, $E[Y(Z,0) \mid S = s]$ and $E[Y(Z,1) \mid S = s]$, at $S = 600$. This exclusion assumption ensures that the impact of D at $S = 600$ is entirely driven by take-up and is unrelated to other factors slightly above or below 600 points, once a student has already taken the scholarship.

The third assumption is that $E[D(0) \mid S = 600] \neq E[D(1) \mid S = 600]$. This first-stage assumption ensures that the denominator of β_{FRD} is not equal to zero.

The final assumption is $\lim_{s \to 600^-} P[D(0) = 1, D(1) = 0 \mid S = s] = \lim_{s \to 600^+} P[D(0) = 1, D(1) = 0 \mid S = s] = 0$. This monotonicity (or no-defier) assumption indicates that no students slightly below 600 points would stop enrolling in the scholarship had they crossed the 600 threshold, a reasonable assumption. Along with the first-stage assumption, this implies that the denominator of β_{FRD} converges to the probability of compliers, which is greater than zero.

I use local linear regressions with a uniform kernel and socioeconomic and demographic characteristics to estimate the means in Equation 2. The controls are included as recommended by Calonico et al. (2019), which is equivalent to a 2SLS strategy with linear controls that are not interacted with the threshold. Robustness checks using alternative bandwidths, kernel functions, polynomial specifications, and specifications without controls are detailed in Online Appendix B.

My main results are estimated using a bandwidth that minimizes the Mean Squared Error (MSE) following the procedure in Calonico et al. (2014) , which builds upon Imbens and Kalyanaraman (2012). I estimate the optimal bandwidth for all of my main outcomes in higher education and teaching markets and use the smallest one (21.9 PSU points) to present results that are consistent in terms of both the sample and the specification. In Online Appendix B, I test the bandwidth sensitivity of my results using the optimal bandwidth for each outcome, along with five other alternatives. For statistical inference, I use non-parametric robust p-values as also proposed by Calonico et al. (2014).

I check whether test retaking behavior could violate the exclusion restriction and find that it is not an issue. Students slightly below the cutoff could retake the test until they cross the 600-point barrier. In this case, the cutoff might not only have induced some students to join the program, but it could also have provided more information to students not crossing the threshold due to test re-taking behavior, including details about the scholarship and academic institutions. Trimming is a common approach to address this issue in related RDD loan program studies (Solis, 2017; Bucarey et al., 2020; Aguirre, 2021) at the cost of reducing external validity. Avoiding having to implement this method is particularly relevant for this case since almost half of the recipients in my sample took the PSU more than once (see Table 2). Table 3 indicates no effect on test retaking, consistent with the results in Table 1, which show no effect of the 600-point offer on program take-up in subsequent years. Table 1 also illustrates how this could be a substantial problem for State Guaranteed Loans, an issue addressed by Solis (2017), Bucarey et al. (2020) and Aguirre (2021).

Table 3: Impact of Grants and Loans on Retaking the Admissions Test

	T+1 Academic Year	T+2 Academic Year	T+3 Academic Year	T+4 Academic Year
BVP (600 points)	-.006	-.003	-.002	.001
	(.005)	(.003)	(.002)	(.002)
Alternative Policies				
HE Grants (550 points)	-.006	-.002	.001	-.001
	(.005)	(.003)	(.002)	(.002)
HE Loans (475 points)	-.055***	-.012***	-.004**	-.005***
	(.003)	(.002)	(.001)	(.001)

Notes. This table shows that the BVP does not have an impact on re-taking the PSU, a major concern that could lead to bias in RDD estimates. I show that this is an issue for State Guarantee Loan impact on re-taking the PSU at the 475 threshold, as illustrated by Solis (2017) and Bucarey et al. (2020). Additionally, I show that this is not a concern for alternative grants: Beca Bicentenario and Beca Juan Gómez Millas which required at least 550 points. I use MINEDUC data the same sharp RDD implemented in the previous table.

In Online Appendix B, I provide additional robustness checks, including a balance test using the socioeconomic and demographic characteristics that support the assumption that significant results are solely due to the treatment (the exclusion restriction). The previous section already depicted a strong first-stage in Figure 2. Appendix A shows that teaching program enrollment is not affected by the threshold in the years prior to the program, as a placebo test, supporting the idea that being above or below the cutoff is unrelated to potential outcomes (the independence assumption). This also suggests that significant results are solely due to the treatment (the exclusion restriction). A similar placebo result is shown in Figure 3 for financial aid, with no changes in take-up of grants and loans for teaching programs around 600 points in the year prior to the implementation of the program. In the same vein, Online Appendix B also displays non-parametric mean plots of all the results for admission scores around the threshold, illustrating that there is a clear discontinuity in the outcomes around 600 points, but not in others within a 100-point range, including the 21.9 bandwidth.

The admission score distribution is not continuous due to the limited number of questions on the test, meaning that the estimator in Equation 2 will not converge to β_{FRD} even with an infinite number of students, unless a parametric extrapolation on the mean estimators is applied. However, this may not be an issue since there are 1,263 mass points, a moderately high number of unique score values. Most importantly, there are multiple values around the cutoff, with several observations as close as 0.005 standard deviations to 600 points (Cattaneo et al., 2024). The latter is clearly visible in Figure 2 and in the score histogram depicted in Online Appendix B. Additionally, results are robust to polynomial specifications, limiting concerns about parametric extrapolation. Nevertheless, I avoid clustering observations at the PSU score level, following

Kolesár and Rothe (2018).

A key implication of the first fuzzy RDD assumption is the absence of manipulation of the admission score. This is unlikely and has never been an issue in the centralized admission test in Chile. Nevertheless, I document in Online Appendix B some level of imbalance in the score density around 600, which also occurs for placebo cutoffs, probably due to the discreteness of the admission scores. Because of this manipulation test rest results, even though manipulation is unlikely, I include a donut hole robustness check by estimating a sample excluding one quarter of the bandwidth around 600 points. My results in Online Appendix B confirm that the imbalance in the score density does not seem to affect the results.

3.2 Method: Complier Decomposition, Characterization and Counterfactuals

3.2.1 Decomposition and Costs

To gain more insight into student choices, take-up, and costs from β_{FRD}, I decompose BVP compliers at 600 into marginal and inframarginal teaching program students. Inframarginal teaching program compliers would have studied teaching even if they scored slightly below 600 points and not taken the BVP, while marginal compliers would not have. I find that 9.2% of test-takers around 600 points were compliers, with one-third classified as marginal and two-thirds as inframarginal enrollees in the teaching program. This implies that recruiting one additional student for a teaching program with 600 points requires three scholarships.

I begin by noting that the individual treatment effect on a binary outcome Y^* can only take three values: $Y_1^* - Y_0^* \in \{-1, 0, 1\}$ where $Y_D \equiv Y(D)$ represents the potential outcome of a student if treated ($D = 1$) or untreated ($D = 0$). Based on these potential outcomes, students are categorized into a set G with three elements: $G \in m, im, rm$ Students can be "marginal" on outcome Y^* ($G = m$), because they would increase the number of observations taking the value one to the sample if treated ($Y_1^* - Y_0^* = 1$), "inframarginal" on outcome Y^* ($G = im$), because they would not change the sample distribution of Y^* if treated ($Y_1^* - Y_0^* = 0$) or "reverse marginal" on outcome Y^* ($G = rm$), because they would decrease the number of observations taking the value one to the sample if treated ($Y_1^* - Y_0^* = -1$). Then, the average treatment effect of a binary variable Y^* equals the probability a student is marginal on Y^* minus the probability a student is reverse marginal on Y^*.

$$E(Y_1^* - Y_0^*) = -1 * P(G = rm) + 0 * P(G = im) + 1 * P(G = m) = P(G = m) - P(G = rm) \tag{3}$$

For some binary outcomes, it may be reasonable to assume that the treatment will never be 'reverse marginal.' In these cases, the treatment effect is monotone on Y^*, meaning it will either maintain or increase this outcome if treated, but never decrease it. This outcome monotonicity assumption applies to the service scholarship program for the outcome of enrolling in a teaching program, as it is impossible to prevent a higher education test-taker to stop enrolling in a teaching program if they take the teaching scholarship.

Additionally, Equation 3 clearly shows that under outcome monotonicity, the average treatment effect equals $P(G = m)$.

If outcome monotonicity on Y^* applies to any admission's test-taker, it also applies to compliers at 600 points. Hence, under outcome monotonicity, and from Equation 1 and Equation 3 , the fuzzy RDD effect of Y^* is equal to the probability that a complier student at 600 points is marginal on outcome Y^*:

$$E(Y_1^* - Y_0^* | S = 600, C) \equiv \beta_{FRD} = P(G = m | S = 600, C) \tag{4}$$

Furthermore, since under outcome monotonicity, $G \in \{m, im\}$, the probability that a complier student at 600 points is an inframarginal on outcome Y^* equals:

$$1 - \beta_{FRD} = P(G = im | S = 600, C) \tag{5}$$

In Table 4, in the higher education results section, I show that the estimate of β_{FRD} for enrolling in a teaching program is 0.33. This means that one-third of the compliers at 600 points are marginal and two-thirds are inframarginal, which implies that enrolling one additional teaching program student at 600 points requires enrolling three students in the scholarship.

3.2.2 Characterizing Compliers

I seek to understand who was induced to take up the program and change their higher education choices at 600 points. I find that compliers have a similar socioeconomic status compared to teaching program students below the cutoff, while the marginal compliers who changed their higher education choices are more likely to be older and male than inframarginals.

Defining X as a pre-treatment characteristic, I estimate $E[X|S = 600, C]$, the average characteristic of the compliers, by replacing Y with XD in the sample analogue of Equation 2 following Pinotti (2017). Similarly, I estimate $E[X|S = 600, C, G = m]$, the average characteristic of the marginal compliers, by replacing Y with XY and D with Y in the sample analogue of Equation 2 , following Comey et al. (2024)[19]. Estimating the latter requires the assumption that $\lim_{s \to 600^+} E[X|S = s] = \lim_{s \to 600^-} E[X|S = s]$, which is tested in the balance test. This assumption also relies on the independence of the characteristic X and a binary instrument, as stated in Proposition 1 of Comey et al. (2024) in the context of a fuzzy RDD. Finally, I use the decomposition to derive Equation 6, where $E[X|S = 600, C, G = im]$ can be estimated by solving this equation, as it is the only missing estimate.

[19] Comey et al. (2024) calls what I define as marginal compliers for a specific outcome as a "supercomplier".

$$E[X|S=600,C] = E[X|S=600,C,G=m]P(G=m|S=600,C)+ \\ E[X|S=600,C,G=im]P(G=im|S=600,C) \tag{6}$$

I compare these results with the mean of students enrolled in a teaching program with slightly fewer than 600 points. Almost none of them took the scholarship[20]. I refer to this group as the 'Always Teaching Program' (ATP) students because crossing the 600-point cutoff would not have diverted them from the teaching programs. It was not possible to take the scholarship without studying teaching, and there were no other cutoffs at 600 points that would have redirected students to another major[21].

In Table 2 from the context section, I show that ATP students came from much lower-income backgrounds compared to the average 600-point student across various metrics. For example, the probability that an ATP student came from a private, unsubsidized school is half that of an average student at 600 points (19.2 percentage points), and similar to the average test-taker (10.8 percentage points). ATP students had a 26 percent higher probability of coming from a lower-income family (42.6 percentage points), a 30 percent lower probability of having private health insurance (25.2 percentage points), and their parents had studied around 0.7 years less (around 12.3 years). ATP outcomes will be used later in the paper to build a counterfactual because they represent teaching program students with no service scholarship incentives[22].

The compliers share many similar characteristics with ATP students. Column (6) in Table 2 shows that the probability of a test-taker scoring 600 points being a complier is 9.2. Compliers were, on average, 0.58 years older than ATPs and 5 percentage points more likely to come from a capital city, while having a very similar socio-economic background. They also had a comparable profile on many other measures, including the probability of being from a lower-income family, parents' education, and school background. Estimates of these differences, along with confidence intervals that account for non-parametric bias, are presented in Table A5 in the Appendix A.3.

Marginal compliers characteristics in column (7) are, in general, quite variable. However, two statistically significant differences from inframarginal students stand out: marginal compliers are 1.92 years older and are 16.6 p.p. more likely to be male compared to inframarginals (see Table A5 in the Appendix A.3).

The profiles of different types of compliers reaffirm the idea that recipients tend to come from more disadvantaged backgrounds than the average 600-point student. As mentioned earlier, this is particularly relevant for the increased likelihood of failing to comply with the agreement and converting the grant into a

[20]The exceptions are "always takers" in column (5) of Table 2 , students who received an offer at 580 point because they were in the top 5% in their high school cohort and are also slightly before 600 points. This group amounts to 0.7% of students at 600 points, and not including them does not substantially change any result in these comparisons or later in the paper.

[21]Figure A2 in Appendix A.2 , shows the 600-point threshold is exclusive to BVP-eligible programs by comparing the distribution of teaching program enrollment for eligible and ineligible institutions before and after the BVP started.

[22]See footnote above.

loan.

3.2.3 Treatment Effect Baseline and Counterfactuals

I construct one baseline and two counterfactuals to assess the practical (or economic) significance of the fuzzy estimates (Lipsey et al., 2012). First, I estimate as a baseline $E[Y(0)|S=600,C]$, the mean potential outcome under no treatment for the compliers at 600 points to compare the treatment effect to a scenario with no treatment. The baseline corresponds to the mean control group outcome baseline in a hypothetical RCT with perfect compliance, where the study sample is already restricted to the compliers at 600 points. I derived this baseline by using Lemma 2.1 of Abadie (2002) . Notably, this approach does not require any additional assumption besides the fuzzy RDD estimates assumptions[23]. The identification of the estimand of interest is presented in Equation 7, which I estimate using the same local linear mean procedure employed for the fuzzy RDD estimates in the main results.

$$\beta_{baseline} = E[Y(0)|S=600,C] = \frac{\lim_{s\to 600^+} E[Y(1-D) \mid S=s] - \lim_{s\to 600^-} E[Y(1-D) \mid S=s]}{\lim_{s\to 600^+} E[(1-D) \mid S=s] - \lim_{s\to 600^-} E[(1-D) \mid S=s]} \tag{7}$$

Second, Fuzzy RDD coefficients related to teaching outcomes alone may not yield easily interpretable estimates in the teaching context. For example, an increase in the probability of dropping out from a teaching program may sound like an undesirable result. However, the policy may bring more students into teaching programs who could drop out at a lower rate than high-scoring teaching program students typically do, while also decreasing the dropout rate among those who would study teaching regardless.

To address this issue, I construct two counterfactuals to evaluate how much the treatment could affect recipients' trajectories beyond merely enrolling additional students. In the first counterfactual, I estimate what the fuzzy estimates would be if they reflected nothing more than enrolling additional teaching program students who follow the same trajectory as those slightly below the cutoff[24]. This counterfactual equals the average outcome of ATP students—teaching program students with 600 points just below the cutoff—multiplied by the probability that a recipient is a marginal complier in teaching program enrollment, as described in Equation 8. If 20% of the teaching program students just below the cutoff drop out of higher education, and 33% of the complier recipients are marginal students, then the effect of enrolling students assuming they make identical choices—without any further incentives—would be $\beta_{TPcntr1} = 0.33 \times 0.2 = 0.066$.

$$\beta_{TPcntr1} = \lim_{s\to 600^-} E(Y|TP=1, S=s)P(G=m|S=600,C) \tag{8}$$

[23] Abadie (2002) method is for Instrumental Variable design (IV), but as pointed out by Angrist and Pischke (2009) fuzzy RDD is IV where the instrument is the eligibility binary indicator Z described on section 3.1.

[24] As those in the teacher pipeline illustrated above in Figure 1 but below the cutoff.

Teaching program students slightly below the cutoff represent a policy-relevant counterfactual because they provide a comparison between the fuzzy estimates and the trajectories of untreated high-scoring teaching program students, as depicted in the context section, but within this sample. However, compliers may have different unobservable characteristics compared to teaching program students slightly below the cutoff[25]. For this reason, in the second counterfactual, I estimate what the fuzzy estimates would be if they reflected nothing more than enrolling additional teaching program students who follow the same trajectory as untreated compliers who would study teaching regardless of the program—the inframarginal compliers. The treatment would only enroll students with the same behavior of teachers who would study regardless to more students. In practice, $\beta_{TPcntr2}$ equals the baseline outcome of inframarginal students following enrollment, which corresponds to $\beta_{baseline}$ from Equation 7, divided by the share of inframarginal students and multiplied by the share of marginal compliers in teaching program enrollment[26].

$$\beta_{TPcntr2} = \frac{\beta_{baseline}}{P(G = im|S = 600, C)} P(G = m|S = 600, C) \tag{9}$$

4 Results

This section is divided into three parts. In the first part, I present results on higher education, including the effects on competing financial aid programs. In the second part, I present teaching market results, including teaching by school type and trajectories over time in relation to the agreement. Finally, in the third part, I provide statistics on indebtedness at the end of the agreement.

4.1 Higher Education Results

Similar to prior research, I find a sizable impact on attracting students with 600 points to teaching programs. However, I also find a large effect on dropping out of or switching from these programs. I show that after enrollment, all the results replicate the low-persistence patterns observed in teaching programs, leading many recipients to violate the original agreement, converting their grants into loans.

Column (2) in Table 4 shows that the program increased the probability of a student enrolling in a teaching program by 33 percentage points from a baseline of 65.6 p.p, but it increased the probability of graduating from a teaching program by only 16.9 percentage points from a baseline of 44.3 p.p. The graduation effect not only appears lower than the enrollment effect but is also 4.9 percentage points lower than the expected graduation rate for teaching program students scoring slightly below 600 points (the ATPs), as presented in column (3). This lower graduation rate is attributed to a 6.4 percentage point increase in switching out of the program and a 9.7 percentage point increase in dropping out of higher education.

[25] Teaching program students below the cutoff are obviously not the same as any test-taker below the cutoff, which is the comparison group in the ITT.

[26] For more detail see proofs in the Online Appendix B.3.

The program increased on-time graduation by 7.7 p.p, which is almost equal to any of the two counterfactuals. This suggests that the program did not accelerate graduation for recipients, despite the threat of losing the scholarship and still having to repay. Although moral hazard due to drop-out insurance implemented at the end of 2017 could also explain these effects, the graduation rates for 2011 recipients and 2013 recipients near the 600-point cutoff are almost identical (63.1 p.p. for 2011 and 62.98 p.p. for 2013 within the bandwidth), despite the latter cohort being exposed to this policy change, which occurred nearly seven years after the first cohort enrolled in the program.

Table 4: Higher Education Effects

	Fuzzy RDD Baseline Mean (1)	Fuzzy RDD Treatment Estimate (2)	ATPs Mean (3)	Immediate TP Counter-factual1 (4)	Immediate TP Counter-factual2 (5)
Panel A - Outcomes Following Immediate Enrollment					
Enrolled in HE	0.873***	0.130*			
	(0.056)	(0.055)			
Enrolled at TP	0.656***	0.330***	1.000***	0.330	0.323
	(0.041)	(0.041)	(0.000)		
Dropped out from TP & Switched	0.095***	0.064**	0.150***	0.05	0.047
	(0.015)	(0.02)	(0.02)		
Dropped out from TP & HE	0.111***	0.097**	0.183***	0.06	0.055
	(0.017)	(0.022)	(0.022)		
Graduated from HE	0.674***	0.109	0.814***		
	(0.073)	(0.073)	(0.022)		
Graduated from TP	0.443***	0.169***	0.661***	0.218	0.218
	(0.031)	(0.036)	(0.026)		
Graduated on time from HE	0.221***	0.032	0.261***		
	(0.052)	(0.054)	(0.023)		
Graduated on time from TP	0.158***	0.077**	0.240***	0.079	0.078
	(0.018)	(0.023)	(0.022)		
Panel B - Outcomes Following Ever Enrollment					
Ever Enrolled in HE	0.987***	0.013			
	(0.017)	(0.017)			
Ever Dropped out from HE	0.186*	-0.012			
	(0.054)	(0.055)			
Ever Graduated from HE	0.775***	0.019			
	(0.061)	(0.061)			
Ever Enrolled at TP	0.658***	0.328***			
	(0.046)	(0.046)			
Ever Dropped out from TP	0.210***	0.140**			
	(0.026)	(0.032)			
Ever Graduated from TP	0.442***	0.184***			
	(0.036)	(0.04)			

Notes. (i) HE: Higher Education, TP: Teaching Program. (ii) (1) Fuzzy RDD Baseline Mean: The average potential outcome of compliers at the cutoff (600 points) if they had not been treated. (2) Fuzzy estimates (3) ATPs Mean (Always Teaching Program students): The mean of students enrolled at a TP below the cutoff, estimated as the mean below the cutoff of a sharp RDD without controls in a sample restricted to teaching program students. (4) Immediate TP Counterfactual 1: This is equal to 0.330, the fuzzy estimate of the effect on the probability of enrolling at a TP, times column (3) for teaching program outcomes following immediate enrollment. The counterfactual shows how would be the fuzzy estimate if the policy only attracted more students similar to teaching program students who immediately enrolled after scoring slightly below 600 points (5) Immediate TP Counterfactual2: This equal to the baseline multiplied by 0.493. The counterfactual shows how would be the fuzzy estimates if the policy did nothing, rather than attracting more students similar to inframarginal students. (iii) All the RDD estimates use a local linear regression, a uniform kernel, and a 21.9 bandwith, the minimum bandwidth that minimizes the MSE following Calonico et al. (2014) across all the outcomes. (iv) Heteroscedasticity robust standard errors in parenthesis. (v) *p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01 of estimates considering non-parametric bias and heteroscedasticity using the robust estimates of Calonico et al. (2014).

The results show changes in trajectories toward teaching, but not toward higher education, except for the timing of enrollment. Panel B of Table 4 shows that the program did not encourage students to ever enroll in higher education. The results are null and precise, from a baseline of 98.7 percent of the test-takers with 600 points enrolled at some point. This provides evidence that there were no short-term credit constraints for enrolling in higher education near the 600 points cutoff, unlike other grants and loans policies in Chile, which require much lower admission scores (Solis, 2017; Bucarey et al., 2020; Aguirre, 2021).

Similarly, the effect of ever graduating from higher education is also null from a 77.5 p.p. baseline, indicating that BVP incentives and the generous grant did not help students enroll or remain in higher education, and dropout rates were already high (18.6 p.p baseline). Results for ever graduating or dropping out of a teaching program differ slightly from those for graduating or dropping out in panel A of Table 4. This difference is partly due to the small impact on take-up on the second year as presented in Table 1 and discussed in the context section.

The effect of enrolling in higher education on the year of admissions is 13 p.p, which accounts for 39.39% of marginal compliers (0.13/0.33). This indicates that these students chose teaching over waiting to retake the test later, consistent with present bias in educational choices (Lavecchia et al., 2016). The effect of ever enrolling in higher education is null, and no other cutoff affects students at 600 points. Thus, this effect impacts only marginal compliers' enrollment because, by definition, inframarginal compliers do not change their enrollment choices.

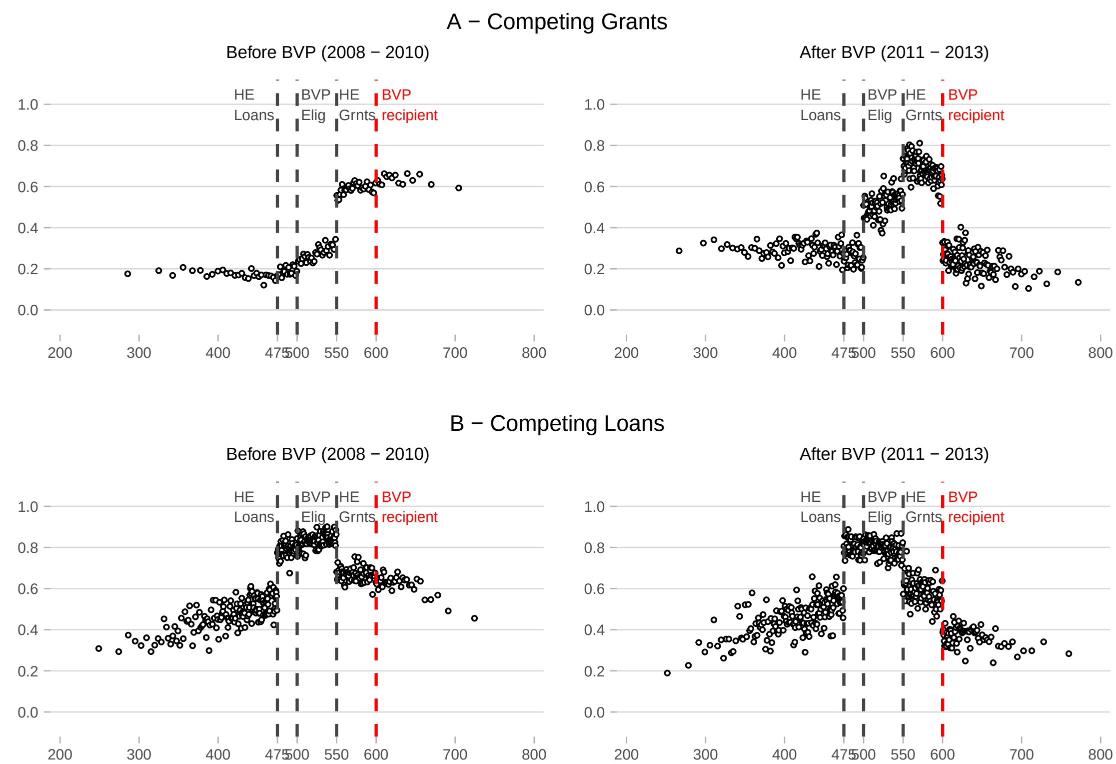
Table 5: Effects on Ever Taking Competing Grants and Loans in Any Year After the Test

	Mean Below 600 Baseline (1)	ITT RDD Estimates (2)	Fuzzy RDD Baseline Mean (3)	Fuzzy RDD Treatment Estimates (4)
Panel A - Grants				
Juan Gomez Millas	0.151***	-0.005	0.084	-0.049
	(0.006)	(0.005)	(0.053)	(0.052)
Bicentenario	0.366***	-0.054***	0.607***	-0.572***
	(0.008)	(0.006)	(0.07)	(0.062)
Other grants	0.330***	-0.027***	0.457***	-0.286***
	(0.008)	(0.006)	(0.069)	(0.065)
Any Grant	0.585***	-0.044***	0.673***	-0.465***
	(0.008)	(0.006)	(0.072)	(0.06)
Panel B - Loans				
Traditional University Loan	0.283***	-0.027***	0.470***	-0.283***
	(0.007)	(0.006)	(0.065)	(0.063)
State Guaranteed Loan	0.419***	-0.006	0.257	-0.059
	(0.008)	(0.007)	(0.072)	(0.073)
Any Loan	0.610***	-0.024	0.590***	-0.258
	(0.008)	(0.007)	(0.071)	(0.069)

Notes. (i) HE: Higher Education, TP: Teaching Program. (ii) (1) Mean Below 600: The mean below the cutoff of the local linear ITT without controls. (2) ITT Estimates: Intent to Treat Estimates, using a sharp RDD with controls. (3) Fuzzy RDD Baseline Mean: The average potential outcome of compliers at the cutoff (600 points) if they had not been treated. (4) Fuzzy estimates (iii) All the RDD estimates use a local linear regression, a uniform kernel, and a 21.9 bandwidth, the minimum bandwidth that minimizes the MSE following Calonico et al. (2014) across all the outcomes. (iv) Heteroscedasticity robust standard errors in parentheses. (v) *p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01 of estimates considering non-parametric bias and heteroscedasticity using the robust estimates of Calonico et al. (2014).

In terms of financial aid, I find that the BVP substantially decreased the probability of ever receiving substantial no-risk grants and ever taking subsidized loan programs. At the time of enrollment, converting the grant into a loan was risky. Since recipients had their tuition fully covered, they could not apply for no-risk grants while receiving the scholarship. As a result, the full grant could convert into a loan much larger than the policy limit for loan aid, which was capped at less than 80% of tuition to prevent over-borrowing (Eluchans, 2014; Beyer et al., 2015). The last section provides further statistics, including those related to the implementation of drop-out insurance in 2017.

Figure 3: Competing Grants and Loans Take-up in Teaching Programs



Notes. (1) These figures illustrate how the BVP significantly crowded out competing publicly funded grants and loans in teaching programs and confirm that the BVP cutoffs did not play any role prior to its implementation. They also highlight the connection between all grants and loans in teaching programs and the substantial probability of keep taking loans in teaching programs above 600 points after 2011. (2) The non-parametric means represent the share of teaching program students who enrolled after admission and took up any grant or loan, other than the BVP, at any time in their higher education trajectory after taking the test for different PSU scores. These means are estimated using the IMSE-optimal quantile-spaced method with polynomial regression, implemented in the rdplot R command as described in Calonico et al. (2015a) and Calonico et al. (2015b). (3) The vertical lines indicate cutoffs at different admission scores for the most popular grants and loans, including the BVP. HE Loans: TUL and SGL loans cutoff (475 points); HE Grnt: BC and JGM grants cutoffs (550 points); BVP Elig: BVP Eligible institutional cutoff (500 points); BVP Recipient: BVP Recipient cutoff (600 points).

In Table 5 I show that the program reduced the take-up of competing grants by 46.5 p.p. Most notably, it decreased the probability of ever receiving the BC grant, the largest grant covering 80% of tuition, by 57.2 p.p. It also reduced the probability of ever taking a higher education loan by 25.8 p.p. Despite the results not being significant using non-parametric robust p-values, I interpret this as a significant result since the estimate is identical and significant at a 1% level of significance when considering non-parametric estimates and an optimal MSE bandwidth, instead of the common bandwidth of 21.9 p.p. The equivalent to Table

5 with MSE bandwidth is presented in the Online Appendix B along with extensive bandwidth stability sensitivity results for all the other outcomes of this paper, for which this is not an issue.

Finally, the strong interaction between different grants and loan programs helps explain why other programs reduced short-run credit constraints, whereas this program did not (Solis, 2017; Bucarey et al., 2020; Aguirre, 2021). Since many students came from lower-income test-taking families, recipients already had access to several financial aid alternatives at 600 points. More than 60% of teaching program students, close to 600 points took either a grant or a loan before the service scholarship was implemented. Figure 3 illustrates the high take-up of competing grants and loans in teaching programs before 2011 and how the program crowded out both of them after its implementation, as discussed above.

Notably, the last graph in Figure 3 reveals that nearly 40% of teaching program students with scores above 600 still ended up taking subsidized loans, even though most of them received the service scholarship, as noted in Figure 2 in Appendix A. This is not as uncommon as one might expect, as only about 20% of teaching program students slightly below the 600-point threshold graduated on time, as shown in the third column of Table 5. Delayed graduation is an annually reported national issue in Chile (Mi Futuro, 2024), affecting even its highest-quality institutions (ProrrectoríaUC, 2023). Although the scholarship was marketed as 'Free Education' (see advertisement in Figure A1), it did not fund students beyond the official program duration. Therefore, additional funding was often needed, both for students who extended their time in teaching programs and for those who switched to other programs.

4.2 Teacher Market Results

I find that the BVP recruited further students to teach in publicly funded schools over time, even after the agreements ended. Importantly, all the results, including dynamic effects and different measures of teaching, align with both of my untreated teaching-program counterfactuals, suggesting that the impact of the scholarship can be entirely attributed to bringing marginal students into teaching, without any effect from the threat of converting the grant into a loan.

The service scholarship brought more students into publicly subsidized schools and had no effect on unsubsidized schools, which goes against the agreement. Panel A of Table 6 shows that the BVP increased the probability that a complier test-taker ended up teaching by 18.1 p.p, which is almost identical to the impact on teaching at a publicly funded school (17.8 p.p). The effect on teaching or only teaching at publicly funded schools is positive but small. However, since two-thirds of the compliers are inframarginal, with very similar characteristics to teaching program students below the cutoff, and around 84% of the program recipients, diverting students from teaching at an unsubsidized school should have had a negative effect. This goes against the agreement, which is confirmed by my baseline estimates where 6.7 p.p kept teaching at a publicly unsubsidized school and post-treatment outcomes in the next section.

Table 6: Teaching Effects Following Immediate Enrollment

	Fuzzy RDD Baseline Mean (1)	Fuzzy RDD Treatment Estimate (2)	ATPs Mean (3)	Immediate TP Counter-factual1 (4)	Immediate TP Counter-factual2 (5)
Panel A - Taught by Type of School					
Taught	0.375***	0.181***	0.549***	0.181	0.185
	(0.028)	(0.033)	(0.027)		
Taught pub. funded	0.339***	0.178***	0.492***	0.163	0.167
	(0.026)	(0.032)	(0.027)		
Taught priv. unsubsidzed	0.067***	0.021	0.102***	0.034	0.033
	(0.012)	(0.015)	(0.016)		
Panel B - Taught in Publicly Funded Schools: Years					
Taught only one year	0.039***	0.021	0.058***	0.019	0.019
	(0.009)	(0.012)	(0.013)		
Taught only two years	0.041**	0.021**	0.059***	0.019	0.02
	(0.008)	(0.012)	(0.012)		
Taught at least three years	0.258***	0.136***	0.376***	0.124	0.127
	(0.022)	(0.029)	(0.026)		
Panel C - Taught in Publicly Funded Schools: Contract Hours					
Employed 132 hours or less	0.138***	0.050**	0.192***	0.063	0.068
	(0.016)	(0.021)	(0.02)		
Employed more than 132 hours	0.221***	0.119***	0.323***	0.107	0.109
	(0.021)	(0.027)	(0.025)		

Notes. (i) HE: Higher Education, TP: Teaching Program. (ii) (1) Fuzzy RDD Baseline Mean: The average potential outcome of compliers at the cutoff (600 points) if they had not been treated. (2) Fuzzy estimates (3) ATPs Mean (Always Teaching Program students): The mean of students enrolled at a TP below the cutoff, estimated as the mean below the cutoff of a sharp RDD without controls in a sample restricted to teaching program students. (4) Immediate TP Counterfactual1: This is equal to 0.330, the fuzzy estimate of the effect on the probability of enrolling at a TP, times column (3) for teaching program outcomes following immediate enrollment. The counterfactual shows how would be the fuzzy estimate if the policy only attracted more students similar to teaching program students who immediately enrolled after scoring slightly below 600 points (5) Immediate TP Counterfactual2: This is equal to the baseline multiplied by 0.493. The counterfactual shows how would be the fuzzy estimates if the policy did nothing, rather than attracting more students similar to inframarginal students. (iii) All the RDD estimates use a local linear regression, a uniform kernel and a 21.9 bandwidth, the minimum bandwidth that minimizes the MSE following Calonico et al. (2014) across all the outcomes. (iv) Heteroscedasticity robust standard errors in parenthesis. (v) *p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01 of estimates considering non-parametric bias and heteroscedasticity using the robust estimates of Calonico et al. (2014).

The program also increased the probability of teaching in publicly funded schools even after the agreement ends, a particularly relevant result since productivity improves during the initial years of teaching (Rivkin et al., 2005; Papay and Kraft, 2015). Similarly to the results for higher education, these trajectories were close to the first counterfactual and reflect what was already happening in the year before the test (see Figure 1), suggesting that the effects were not due to the agreement or a short-term income increase beyond increasing enrollment in teacher programs. Comparing fuzzy estimates with the second counterfactual that expands untreated infra-marginal students drives to exactly the same conclusion. Panel B of Table 6 shows null or much smaller results for teaching one or two years in publicly funded schools (around 2 p.p.) compared to teaching for at least three years (13.6 p.p). On average, there was an increase of 0.72 year within 12 years after enrolling in the program.

Finally, the effect on teaching in publicly funded schools over time also replicates in terms of contract hours. Contract hours are particularly relevant because the requirements became more flexible in January 2017, with the specification that the requirements would be measured with contract hours. Having been employed for 132 hours -where a one-year contract is 44 hours- can exceed the agreement after the policy change for some types of teachers, as described in the context section. Nonetheless, a recipient who taught more than 132 hours meets any of the updated agreement requirements within 7 years after graduation, which is almost always more than 12 years after enrolling in the scholarship. The effects on contract hours, presented in Panel C of Table 6 replicate the results of years of teaching in Panel B. Most importantly, the program increased the probability that a test-taker at 600 points taught more than 132 hours by 11.9 p.p., working more than required and more than working 132 hours or less. As in all the other panels in the table, results are close to both counterfactuals, showing that the impact of the scholarship can almost entirely be attributed to bringing marginal students into teaching rather than the loan threat or the income increase provided by program.

4.3 Indebtedness Results

In this final subsection, I study the impact of the policy on indebtedness 12 years after taking the scholarship across different scenarios that reflect the policy changes implemented after 2013. Despite adjustments in the policy that prevented nearly 95% of recipients from converting the entire grant into a loan and an increase of 30 p.p. in indebtedness, more than 55% percent still accumulated some form of debt. This outcome ultimately fell short of the 'Studying for Free' promise in the original advertisement and aligns with the financial trajectories of high-scoring students in teaching programs before the program began, as depicted in Figure 1.

The original agreement was extremely difficult for program recipients to comply with, as expected from the teaching trajectories the year before the program's inception (see Figure 1). Column (2) in Table 7shows that, without any policy change, 95% of recipients would have converted their grant into a BVP loan and the BVP would have led to a 35 p.p. increase in indebtedness. Column (4) shows that the introduction of drop-out insurance and partial, gradual repayment options reduced the impact of indebtedness to a negative, non-significant level, relative to a baseline of 60 percentage points, which still fell short of the "Studying for Free" promise.

Table 8 shows the descriptive indebtedness status of the three cohorts 12 years after starting the program. Mirroring the fuzzy RDD results that estimates the effects at the 600 points, nearly 5% of all BVP recipients would have failed to comply with their agreement and would have turned the full grant into a loan in any of the three cohorts examined. In addition to this difficulty, 35% percent of the recipient took additional subsidized loans from the TUL or the SGL programs, which would have made repayment even more difficult in a context where nearly half of the students failed to repay their SGL loans (Guastavino and Miranda, 2019), and around 40% of recipients came from lower-income families (see Table 2).

Table 7: Effects on Indebtedness

	Original Agreement (OA)		OA + Drop-out Insurance + Partial and Gradual Repayment	
	Fuzzy RDD Baseline Mean (1)	Fuzzy RDD Treatment Estimates (2)	Potential No Treatment Mean Baseline (3)	Fuzzy RDD Treatment Estimates (4)
BVP Loan	0	0.953***	0	0.309***
	(0.00)	(0.008)	(0.00)	(0.016)
Additional Loans	0.594***	-0.272	0.594***	-0.272
	(0.075)	(0.073)	(0.075)	(0.073)
Any Loan	0.594***	0.346***	0.594***	-0.065
	(0.075)	(0.072)	(0.075)	(0.074)

Notes. (i) This table shows that the BVP would have led to a 35 percentage point increase in indebtedness without any change in the original agreement. 95 percent of the recipients would have converted their grant into a loan without any policy change. The introduction of drop-out insurance and partial, gradual repayment options reduced the impact of indebtedness to a negative, non-significant level, relative to a baseline of 60 percentage points, which still fell short of the "Studying for Free" promise. (ii) Potential No Treatment Mean Baseline: The average potential outcome of compliers at the cutoff (600 points) if they had not been treated. (iii) All the RDD estimates use a local linear regression, a uniform kernel, and a 21.9 bandwidth, the minimum bandwidth that minimizes the MSE following Calonico et al. (2014) across all the outcomes. (iv) Heteroscedasticity robust standard errors in parenthesis. (v) *p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01 of estimates considering non-parametric bias and heteroscedasticity using the robust estimates of Calonico et al. (2014). The baseline estimate of taking the BVP is equal to a non-random zero.

Table 8: BVP Recipients Indebtedness

Cohort	BVP Loan Status					Additional Loans (6)	BVP or Other. Loans (7)
	Repayed (No Loan) (1)	Insurance (No Loan) (2)	Some Loan (3)	Full Loan (4)	BVP Loan (5)		
Panel A - Original Agreement							
2011	6.07%	0%	0%	93.93%	93.93%	41.59%	96.34%
2012	5.09%	0%	0%	94.91%	94.91%	33.09%	95.91%
2013	4.38%	0%	0%	95.62%	95.62%	29.34%	96.6%
All	5.27%	0%	0%	94.73%	94.73%	35.36%	96.28%
Panel B - Original Agreement plus Drop-out Insurance, Gradual and Partial Repayment							
2011	36.5%	33.37%	17.66%	12.47%	30.13%	41.59%	59.88%
2012	33.65%	34.18%	18.25%	13.92%	32.17%	33.09%	56.2%
2013	30.5%	35.73%	19.9%	13.86%	33.77%	29.34%	55.41%
All	33.87%	34.3%	18.49%	13.33%	31.83%	35.36%	57.42%
Panel C - Current Agreement: Program Data Limited to Graduates Before 2016 or Drop-outs							
2011	28.09%	42.18%	4.37%	25.35%	29.72%	35.29%	58.8%

Notes. This table presents student indebtedness under two hypothetical scenarios and with program data. Panel A calculates how many students would have turned the full grant into a loan under the agreement signed by the three cohorts. Panel B calculates the same but including drop-out insurance (an exemption from repayment when dropping out), gradual repayment (allowing for smaller, but manageable "time" installments instead of having to pay with three years of working teaching hours a week), and partial repayment (allowing repayment of only the unmet portion of obligations). Panel C incorporates all the changes by using program data.

Among the many policy changes, I find that drop-out insurance, gradual repayment, and partial repayment[27] seem to have made dramatic changes to the financial impact of the scholarship. Panel B of Table 8 presents a hypothetical scenario in which only drop-out insurance, gradual, and partial repayment were implemented.

[27] **Gradual repayment:** As described in the Policy Context section, Gradual repayment means the service could be fulfilled by working the equivalent hours of three years in a 44-hour-per-week full-time contract. The scholarship could be repaid in smaller, manageable "time" installments instead than having to pay with three years of working teaching hours a week. **Partial repayment:** Allowing the possibility to repay only the unmet portion of obligations. See context section.

Comparing column (1) in Panel B of Table 8 with Panel A shows that gradual repayment reduced the share of recipients who had to repay the full loan by 28.6 p.p. The same comparison in column (2) shows that drop-out insurance reduced the share of recipients who had to repay the full loan by nearly 35 p.p. A comparison in column (3) shows that partial repayment allowed 18.49% of the recipients to pay only part of the grant instead of the full amount. Finally, the comparison in column (5) shows that the loan was reduced by close to 60 p.p., without considering that nearly two-thirds of the remaining debt is partially repaid with a more flexible scheme. Despite this substantial reduction in loans, column (7) shows that nearly 60% of the recipients would still have some type of loan at the end of 12 years.

Considering all the policy updates and using program data, it is still true that more than half of the recipients ended up with some type of debt. Panel C in Table 8 uses program data, including all the policy updates. This analysis is restricted to the 2011 cohort that graduated before 2016 or dropped out from higher education by 2024, which account for 76% of the recipients. I don't consider the remaining 24% because I have program data only until 2023, and those graduates have not yet completed their updated seven-years retribution period after graduation. However, for the 76% of the 2011 cohort that have already finished their agreement under the new rule, insurance and gradual repayment plays a substantial role. Importantly, more than half still had to take some time of loan.

5 Conclusion

Service scholarships are popular financial aid programs aimed at addressing teacher shortages in areas in-need. Despite their popularity, identifying the causal effect of these programs on recipients, both within and beyond the agreement is challenging. This requires following comparable higher education admission test-takers over a long period of time. This article presented causal evidence showing that a service scholarship program in Chile, which covers full tuition and target high-scoring students to teach in publicly funded schools, can successfully recruit students to enroll, graduate, and teach in publicly funded schools even after the agreement ends, while also driving a significant proportion of students into trajectories outside of the agreement and unmanageable levels of debt. I used national data that include higher education admissions, graduation, financial aid, and teacher-market outcomes for up to 12 years after taking a government service scholarship in Chile. My empirical strategy exploits the similarity between students around the admission requirement threshold to address the selection bias, along with the fact that recipients must enroll in a teaching program to estimate marginal and inframarginal recipients and interpret the causal effects with two counterfactuals.

My results show that the program induced 9.2% of high-scoring admissions test-takers to enroll in the scholarship, with one-third of these changing their career choice and two-thirds studying teaching regardless of the program. Importantly, I show that the recipients signed an agreement that was extremely hard to fulfill without further adjustments. I also find that the program primarily increased enrollment in teaching programs while reproducing low-persistence rates into teaching. This trajectory involved driving several teaching majors out of the program, out of higher education, or out of teaching in publicly funded schools after graduation while at the same time increasing the likelihood that recipients taught beyond the required service period by 11.9 percentage points—a noteworthy outcome given the productivity gains typically observed in the early years of teaching. Although program changes prevented a large share of recipients from receiving substantial no-risk grants, close to 95% failed to meet the original requirements, as anticipated based on previous teaching program trends. Key policy adjustments played a crucial role in preventing most grants from converting into loans. These included making the agreement more flexible through gradual and partial repayment options and introducing dropout insurance seven years after the program's inception. Despite these changes, more than half of the recipients ended up with some form of debt, far from the initial 'Study for Free' promise.

My findings are important not only because they contribute to the limited causal evidence on the teaching recruitment effects under a service agreement over time, and student exposure to potentially difficult-to-repay loans from a widely implemented program addressing a global and recurrent issue, but also because they shed light on four potentially useful policy recommendations for program design and evaluation. First, since I show that most recipients may be driven toward a sizable loan trajectory, my evidence highlights the need for these programs to integrate existing theory and evidence from loan aid programs to improve welfare outcomes. This could begin by applying optimal credit contract theory, which suggests that income-contingent repayment may be essential for enhancing welfare in financial aid programs (Lochner and Monge-Naranjo, 2016; Dearden and Nascimento, 2019). Additionally, loan caps, as discussed in Chile and the US, could help prevent unmanageable debts (Beyer et al., 2015), and there is also emerging evidence on the role of institutional incentives in reducing dropout rates among loan recipients, as demonstrated in Brazil (Barahona et al., 2023). These considerations are particularly relevant in the US, where such policies have a long-standing tradition, but often neglect these aspects (see Program Directory in the Appendix A.1 and Darling-Hammond et al. (2023) and García et al. (2023)). For example, approximately half of TEACH Grant recipients had their grants converted into loans, often exceeding federal loan limits (Barkowski et al., 2018).

Second, the policy design should consider not preventing students from accessing other sources of financial aid while enrolling in the scholarship program. As documented in this work, a large share of recipients may come from lower-income families and opt-out of sizable no-risk grants in exchange for high-risk sizable loans. In any event, the evaluation of the service scholarship should consider examining the crowding out of other financial aid programs when accounting for the total cost of the program (Ramaciotti et al., 2020).

Third, providing clearer and more detailed information about the likelihood of the grant converting into a loan and its repayment conditions may be essential. My findings align with a TEACH grant study showing that recipients may be overly optimistic about fulfilling a service scholarship agreement (Barkowski et al., 2018). Specifically, I showed that the BVP significantly attracted high-scoring students to teaching through an agreement that, at the time, was nearly impossible to fulfill. Consistent with this, in the implementation of the BVP, the policy was advertised as “Free Education” (see Figure A1 in Appendix A.2). However, a few months after signing the agreement, a group of recipients from the first cohort protested that the program was not a “scholarship but a loan” (BioBioChile, 2011). As shown in the last section, before any policy changes by 2013, this claim ultimately proved accurate. A related fact is that the Ministry’s of Education website defines scholarship as follows: “Unlike a loan, you do not need to repay these funds upon graduation or obtaining your degree.” (MINEDUC, 2017). The same issue arises in the US, where many of these programs have been labeled as scholarships (see Program Directory in the Appendix A.1), yet the US Department of Education states: ‘Scholarships are gifts. They don’t need to be repaid’ (FederalAid, 2024). In the same line, many of these programs have been in operation for years or even decades, and information about past recipients’ repayment outcomes is often readily available but rarely disclosed (see the last two columns of the Program Directory in the Appendix A.1). There are, however, some notable exemplary exceptions. For example, the Minority Teacher Education Scholarship (MTES) in Florida (FFMT, 2024) and the Hoosier Educators Scholarship in Indiana (Indiana Commission for Higher Education, 2022) provide information about how likely are past recipients to end up teaching. The Financial Aid Department at Teachers College, Columbia University, warns students about poor repayment outcomes statistics in the TEACH grant (TC, 2024). Addressing this concern is particularly relevant for policies like the Free Teaching Education Scholarship (FTE) in China, where failure to comply with the agreement not only converts the grant into a loan but also imposes a 50% fine, payable in a single installment (Zheng and Shi, 2024).

Finally, I show that, if available, administrative data combined with a causal framework can be used to distinguish and characterize marginal and inframarginal enrollment, providing valuable information for policy evaluation of teacher recruitment. Traditionally, the analysis of this policy has relied on hard-to-obtain and potentially biased surveys to estimate the share of recipients who are marginal and those who would have studied teaching regardless, an area of inquiry dating back to the 1960s[28] (Rogers, 2009; Podolski and Kini, 2016; Valenzuela, 2020). In contrast, this framework leverages existing administrative data that was readily available in the first years of the program, although it may focus on students close to a specific score. For this policy, the Ministry of Finance hired a panel of experts to evaluate the program. The panel conducted a survey asking this question but ended up with a response rate below 16% (Valenzuela, 2020).

[28]“Federal staff worried that the ‘goal of attracting primarily those who would not otherwise have gone into teaching has not so far been attempted’” Rogers (2009) quote of Lawrence and Austin 1967’s Report on the Teach Corps between 1966 and 1967.

References

Abadie A. 2002. Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American Statistical Association* **97**(457): 284–292.

Abraham KG, Kearney MS. 2020. Explaining the Decline in the US Employment-to-Population Ratio: A Review of the Evidence. *Journal of Economic Literature* **58**(3): 585–643. Available from: https://ideas.repec.org/a/aea/jeclit/v58y2020i3p585-643.html.

Aguirre J. 2021. Long-term effects of grants and loans for vocational education. *Journal of Public Economics* **204**: 104539.

Ajzenman N, Bertoni E, Elacqua G, Marotta L, Vargas CM. 2024. Altruism or money? Reducing teacher sorting using behavioral strategies in peru. *Journal of Labor Economics* **42**(4): 000–000.

Angrist JD, Pischke J-S. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Araujo MC, Carneiro P, Cruz-Aguayo Y, Schady N. 2016. Teacher Quality and Learning Outcomes in Kindergarten *. *The Quarterly Journal of Economics* **131**(3): 1415–1453. Available from: https://doi.org/10.1093/qje/qjw016.

Arteaga F, Kapor AJ, Neilson CA, Zimmerman SD. 2022. Smart matching platforms and heterogeneous beliefs in centralized school choice. *The Quarterly Journal of Economics* **137**(3): 1791–1848.

Barahona N, Dobbin C, Ho H, Otero S, Yannelis C. 2023. Skin in the game: Colleges' financial incentives and student outcomes. *Occasional Paper. Retrieved from Https://Theb. Stanford. Edu/~ Sotero/Papers/Skiningame. Pdf*.

Barkowski E, Nielsen E, Noel H, Dodson M, Sonnenfeld K, Ye C, DeMonte E, Monahan B, Eccleston M. 2018. Study of the teacher education assistance for college and higher education (TEACH) grant program. *Office of Planning, Evaluation and Policy Development, US Department of Education*.

Bettinger EP, Long BT, Oreopoulos P, Sanbonmatsu L. 2012. The role of application assistance and information in college decisions: Results from the h&r block FAFSA experiment. *The Quarterly Journal of Economics* **127**(3): 1205–1242.

Beyer H, Alvarado J, Aylwin M, Brunner JJ, Krebs A, Matte P, Molina S, Pavez J, Romaguera P, Rosso PP, Zalaquett P. 2010. Informe final: Primera etapa. Propuestas para fortalecer la profesión docente en el sistema escolar chileno.

Beyer H, Hastings J, Neilson C, Zimmerman S. 2015. Connecting student loans to labor market outcomes: Policy lessons from chile. *American Economic Review* **105**(5): 508–513.

BioBioChile. 2011. Estudiantes de pedagogía denuncian falsa gratuidad en beca vocación de profesor. Available from: https://www.biobiochile.cl/noticias/2011/11/24/estudiantes-de-pedagogia-denuncian-falsa-gratuidad-en-beca-vocacion-de-profesor.shtml.

Bobba M, Ederer T, Leon-Ciliotta G, Neilson C, Nieddu MG. 2021. Teacher compensation and structural inequality: Evidence from centralized teacher school choice in perú. National Bureau of Economic

Research.

Bucarey A, Contreras D, Muñoz P. 2020. Labor market returns to student loans for university: Evidence from chile. *Journal of Labor Economics* **38**(4): 959–1007.

Caetano G, Palacios M, Patrinos HA. 2019. Measuring aversion to debt: An experiment among student loan candidates. *Journal of Family and Economic Issues* **40**: 117–131.

Calonico S, Cattaneo MD, Farrell MH. 2018. On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association* **113**(522): 767–779.

Calonico S, Cattaneo MD, Farrell MH, Titiunik R. 2019. Regression discontinuity designs using covariates. *Review of Economics and Statistics* **101**(3): 442–451.

Calonico S, Cattaneo MD, Titiunik R. 2015a. Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association* **110**(512): 1753–1769.

Calonico S, Cattaneo MD, Titiunik R. 2015b. Rdrobust: An r package for robust nonparametric inference in regression-discontinuity designs. *R J.* **7**(1): 38.

Calonico S, Cattaneo MD, Titiunik R. 2014. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica* **82**(6): 2295–2326. Available from: https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA11757.

Carneiro P, Heckman JJ. 2002. The evidence on credit constraints in post-secondary schooling. *The Economic Journal* **112**(482): 705–734.

Castro-Zarzur R, Espinoza R, Sarzosa M. 2022. Unintended consequences of free college: Self-selection into the teaching profession. *Economics of Education Review* **89**: 102260.

Cattaneo MD, Idrobo N, Titiunik R. 2024. *A Practical Introduction to Regression Discontinuity Designs: Extensions*. Cambridge University Press.

Cattaneo MD, Jansson M, Ma X. 2018. Manipulation testing based on density discontinuity. *The Stata Journal* **18**(1): 234–261.

Chay KY, Greenstone M. 2003. The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *The Quarterly Journal of Economics* **118**(3): 1121–1167.

Chetty R, Friedman JN, Rockoff JE. 2014. Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review* **104**(9): 2633–79.

Claro F, Bennett M, Paredes RD, Wilson T, others. 2013. Incentivos para estudiar pedagogia: El caso de la beca vocacion de profesor. *Estudios Publicos* (131).

Clotfelter C, Glennie E, Ladd H, Vigdor J. 2008. Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in north carolina. *Journal of Public Economics* **92**(5-6): 1352–1370.

Comey ML, Eng AR, Leung P, Pei Z. 2024. Supercompliers. *arXiv Preprint arXiv:2212.14105*. Available from: https://arxiv.org/abs/2212.14105v3.

Currie J, Zivin JG, Mullins J, Neidell M. 2014. What do we know about short-and long-term effects of early-life exposure to pollution? *Annu. Rev. Resour. Econ.* **6**(1): 217–247.

Darling-Hammond L, DiNapoli Jr M, Kini T. 2023. The federal role in ending teacher shortages. *Learning Policy Institute*.

Dearden L, Nascimento PM. 2019. Modelling alternative student loan schemes for brazil. *Economics of Education Review* **71**: 83–94.

Dynarski S, Page L, Scott-Clayton J. 2023. College costs, financial aid, and student decisions, In *Handbook of the Economics of Education*, Elsevier; 227–285.

Ebenstein A, Lavy V, Roth S. 2016. The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics* **8**(4): 36–65.

Einav L, Finkelstein A, Oostrom T, Ostriker A, Williams H. 2020. Screening and selection: The case of mammograms. *American Economic Review* **110**(12): 3836–3870.

Elacqua G, Hincapie D, Hincapié I, Montalva V. 2022. Can financial incentives help disadvantaged schools to attract and retain high-performing teachers? Evidence from chile. *Journal of Policy Analysis and Management* **41**(2): 603–631.

Elacqua G, Hincapié D, Vegas E, Alfonso M, Montalva V, Paredes D. 2018. *Profesión: Profesor En América Latina:¿ Por Qué Se Perdió El Prestigio Docente y cómo Recuperarlo?* Inter-American Development Bank.

Eluchans J. 2014. Modificación de los aranceles de referencia: Nuevo factor de pertinencia laboral. *Serie Evidencias, Centro de Estudios MINEDUC*.

Evans BJ, Boatman A, Soliz A. 2019. Framing and labeling effects in preferences for borrowing for college: An experimental analysis. *Research in Higher Education* **60**: 438–457.

FederalAid. 2024. Finding and applying for scholarships. Available from: https://studentaid.gov/understand-aid/types/scholarships.

Feng L, Sass TR. 2018. The impact of incentives to recruit and retain teachers in "hard-to-staff" subjects. *Journal of Policy Analysis and Management* **37**(1): 112–135.

FFMT. 2024. Florida Fund for Minority Teachers: Scholarship Information. Available from: https://www.ffmt.org/index.cfm?e=inner2&itemcategory=93703.

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): 1–21.

García E, Wei W, Patrick SK, Leung-Gagné M, DiNapoli Jr MA. 2023. In debt: Student loan burdens among teachers. *Learning Policy Institute*.

Glazerman S, Protik A, Teh B, Bruch J, Max J. 2013. Transfer incentives for high-performing teachers: Final results from a multisite randomized experiment. NCEE 2014-4003. *National Center for Education Evaluation and Regional Assistance*.

Guastavino C, Miranda Á. 2019. Determinantes de la morosidad en el pago del crédito con aval del estado.

Available from: https://www.dipres.cl.

Han L, Xie J. 2020. Can conditional grants attract better students? Evidence from chinese teachers' colleges. *Economics of Education Review* **78**: 102034.

He Y, Bartalotti O. 2020. Wild bootstrap for fuzzy regression discontinuity designs: Obtaining robust bias-corrected confidence intervals. *The Econometrics Journal* **23**(2): 211–231.

Imbens G, Kalyanaraman K. 2012. Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *The Review of Economic Studies* **79**(3): 933–959. Available from: https://doi.org/10.1093/restud/rdr043.

Indiana Commission for Higher Education. 2022. Next Generation Hoosier Educators Scholarship Application Open. Available from: https://www.in.gov/che/files/221018_RELEASE_Next-Gen-Scholarship-Application-Open.pdf.

Jackson CK. 2018. What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy* **126**(5): 2072–2107.

Kolesár M, Rothe C. 2018. Inference in regression discontinuity designs with a discrete running variable. *American Economic Review* **108**(8): 2277–2304. Available from: http://www.aeaweb.org/articles?id=10.1257/aer.20160945.

Lavecchia AM, Liu H, Oreopoulos P. 2016. Behavioral economics of education: Progress and possibilities, In *Handbook of the Economics of Education*, Elsevier; 1–74.

Lipsey MW, Puzio K, Yun C, Hebert MA, Steinka-Fry K, Cole MW, Roberts M, Anthony KS, Busick MD. 2012. Translating the statistical representation of the effects of education interventions into more readily interpretable forms. *National Center for Special Education Research*.

Lochner L, Monge-Naranjo A. 2016. Student loans and repayment: Theory, evidence, and policy, In *Handbook of the Economics of Education*, Elsevier; 397–478.

Maestas N, Mullen KJ, Strand A. 2013. Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt. *American Economic Review* **103**(5): 1797–1829.

Manski CF. 1997. Monotone treatment response. *Econometrica: Journal of the Econometric Society*: 1311–1334.

Mi Futuro. 2024. Informes de duración real y sobreduración. Available from: https://www.mifuturo.cl/informes-de-duracion-real-y-sobreduracion/.

MINEDUC. 2017. Becas de arancel. Available from: https://portal.beneficiosestudiantiles.cl/becas/becas-de-arancel.

NASSGAP. 2024. NASSGAP annual survey report on state-sponsored student financial aid. Available from: https://www.nassgapsurvey.com/.

NCES. 2023. Teacher openings in elementary and secondary schools. U.S. Department of Education, Institute of Education Sciences; Condition of Education. Available from: https://nces.ed.gov/programs/coe/indicator/tls.

Neilson C, Gallegos S, Calle F, Karnani M. 2022. Screening and Recruiting Talent at Teacher Colleges Using

Pre-College Academic Achievement. Human Capital; Economic Opportunity Working Group. Available from: https://ideas.repec.org/p/hka/wpaper/2022-004.html.

OECD. 2020. *PISA 2018 Results (Volume v)*. Available from: https://www.oecd-ilibrary.org/content/publication/ca768d40-en.

OECD. 2018. The future of education and skills, education 2030. Available from: https://www.oecd.org/education/2030/E2030%20Position%20Paper%20(05.04.2018).pdf [Accessed 26 November 2019].

Papay JP, Kraft MA. 2015. Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. *Journal of Public Economics* **130**: 105–119.

Papay JP, West MR, Fullerton JB, Kane TJ. 2012. Does an urban teacher residency increase student achievement? Early evidence from boston. *Educational Evaluation and Policy Analysis* **34**(4): 413–434.

Peyton DJ, Dijk W van, Mason-Williams L. 2022. Meeting the moment: Impact of TEACH grant on US undergraduate education degree completion in high-need content areas. *Higher Education Policy*: 1–26.

Pinotti P. 2017. Clicking on heaven's door: The effect of immigrant legalization on crime. *American Economic Review* **107**(1): 138–168.

Podolski A, Kini T. 2016. How effective are loan forgiveness and service scholarships for recruiting teachers? *Learning Policy Institute, Policy Brief*. Available from: https://learningpolicyinstitute.org/product/how-effective-are-loan-forgiveness-and-service-scholarships-recruiting-teachers.

ProrrectoríaUC D de AIyP. 2023. Tasas de titulación y duración real de las carreras: Análisis de las cohortes 2000-2016. Pontificia Universidad Católica de Chile. Available from: https://sustentable.uc.cl/wp-content/uploads/2023/11/Informe-Titulacion-y-Duracion-de-Carreras-cohortes-2000-2016-total-UC.pdf.

Ramaciotti L, Ansoleaga A, Valdebenito MJ. 2020. Informe final de evaluación, evaluación programas gubernamentales programa beca arancel vocación de profesor. External Panel Reprot Finance by Ministry of Education Report. Available from: http://www.dipres.cl/597/articles-205705_informe_final.pdf.

Rivkin SG, Hanushek EA, Kain JF. 2005. Teachers, schools, and academic achievement. *Econometrica* **73**(2): 417–458.

Rogers B. 2009. "Better" people, better teaching: The vision of the national teacher corps, 1965–1968. *History of Education Quarterly* **49**(3): 347–372.

Solis A. 2017. Credit access and college enrollment. *Journal of Political Economy* **125**(2): 562–622.

Steele JL, Murnane RJ, Willett JB. 2010. Do financial incentives help low-performing schools attract and keep academically talented teachers? Evidence from california. *Journal of Policy Analysis and Management* **29**(3): 451–478. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/pam.20505.

TC CU. 2024. Special grant & loan programs.

UNESCO. 2024. Global report on teachers: Addressing teacher shortages and transforming the profession. UNESCO: Paris.

Valenzuela C. 2020. Cambios de comportamiento en ingreso a carreras de pedagogía, e inserción laboral en el sistema escolar con financiamiento estatal. External Panel Reprot Financed by DIPRES, Chile's

Ministry of Finance. Available in the Appendix of Ramaciotti Et Al (2020). Available from: http://www.dipres.cl/597/articles-205705_informe_final.pdf.

Wiederspan M. 2018. Understanding state loan forgiveness and conditional grant programs. MHEC policy brief. *Midwestern Higher Education Compact*.

Zheng Q, Shi Y. 2024. Can service scholarships boost academic performance? Causal evidence from china's free teacher education scholarship. *Higher Education*: 1–25.

A Appendix

A.1 Directory of Service Scholarship Program

Since I could not find a paper or a data-set listing service scholarship for teachers, I built a directory of service scholarship in Tables A1, A2 and Table A3 . I checked all the program websites in Latin America including conditional aid in Elacqua et al. (2018), all the department of education in the US, the programs in the NASSGAP (2024) survey data-set, English speaking countries departments of education in Africa and Asia, and some programs no longer running, including programs discussed in Podolski and Kini (2016) and Wiederspan (2018) and backward citation in some of these papers.

The main goal is not only to assess how popular these programs are, but also to check whether the programs provide information on how likely the recipients are to turn the grant into a loan (Table A1 and Table A3).

Table A1: Service Scholarship Examples Outside of the US (2002 - 2024)

Country	Subregion	Name of the Scholarship	Year	Repayment Details	Repayment Probability
Argentina		Beca Compromiso Docente	2017 - Ongoing	Yes	No
Australia	New South Whales	Teacher Education Scholarships	2002* - Ongoing	Yes	No
Chile		Beca Vocación de Profesor	2011 - Ongoing	Yes	No
China		Free Teacher Education Policy	2007 - Ongoing	No Info	No Info
Nigeria	Bauchi, Katsina, Niger, Sokoto and Zamfara State	Female Teacher Trainee Scholarship Scheme	2008 - 2015	No Info	No Info
Peru		Beca Vocación de Maestro	2016 - Ongoing	Yes	No
South Africa		The Funza Lushaka Bursary Programme	2007 - Ongoing	Yes	No

Notes. *: Could be previous to this date. No Info: Could not find any information.

Table A2: US Service Scholarship Examples No Longer Running (1986 - 2023)

Subregion	Name of the Scholarship	Year	References
Federal level	Paul Douglas Teacher Scholarship	1986 - 1996	Zota (2009)
Delaware	Delaware Teacher Corps	2017-2023	NASSGAP (2024)
Delaware	Christa McAuliffe Incentive Program	Not Found - 2023	NASSGAP(2024)
Georgia	HOPE Teacher Scholarship	1995 - 2013	NASSGAP (2024), Georgia(2024)
Massachusetts	Massachusetts Signing Bonus Program	1999 - 2003	Fowler (2008)
North Carolina	North Carolina Teaching Fellows Program	1986 - 2015	Podolski and Kini (2016)
Washington	Future Teacher Conditional Scholarship	1989 - 1996	Washington(2004)
Maryland	Maryland Teacher Scholarships	1999 - 2014	Maryland (2014)

Table A3: 28 Ongoing US Service Scholarship

Subregion	Name of the scholarship	Year	Repayment Details	Repayment Probability
Federal level	NOYCE Teacher Scholarship Program	2002 - Ongoing	Yes	No
Federal level	TEACH Grant	2007 - Ongoing	Yes	No
Federal level	Woodrow Wilson Teaching Fellowships	2007 - Ongoing	No	No
Arizona	Arizona Teachers Academy	2017 - Ongoing	Yes	No
Arkansas	Arkansas Geographical Critical Needs	2001 - Ongoing	Yes	No
California	Golden State Teacher's Grant Program	2020 - Ongoing	Yes	No
Florida	The Minority Teacher Education Scholarship	1996 - Ongoing	Yes	Yes
Illinois	Golden Apple	1994 - Ongoing	Barely	No
Illinois	Minority Teachers of Illinois (MTI) Scholarship Program	Not Found - Ongoing	Yes	No
Indiana	Next Generation Hoosier Educators Scholarship	2017 - Ongoing	Barely	No
Kansas	Tomorrow's Teacher Scholarship	Not Found - Ongoing	No	No
Kentucky	Teacher Scholarship Program	Not Found - Ongoing	No	No
Maryland	Teaching Fellows for Maryland Scholarships	2014 - Ongoing	No	No
Massachusetts	Tomorrow's Teachers Scholarship	2023 - Ongoing	Yes	No
Massachusetts	MassTeach	2019 - Ongoing	Barely	No
Misourri	Minority Teaching Scholarship	Not Found - Ongoing	No	No
New York	Jose P. Graduate Scholarship Program (Teach NYC)	Not Found - Ongoing	Yes	No
New York	NYS Masters-in-Education Teacher Incentive Scholarship Program	2016* - Ongoing	Yes	No
New York	NYS Math and Science Teacher Incentive Scholarship	Not Found - Ongoing	Yes	No
South Carolina	Teaching Fellowship Program	1999 - Ongoing	Yes	Yes
South Dakota	Critical Teaching Needs Scholarship	2012 - Ongoing	No	No
Tennessee	Tennessee Teaching Scholars Program	1995 - Ongoing	No	No
Texas	Teach for Texas Alternative Certification Conditional Grant	2001 - Ongoing	No	No
Virginia	Virginia Teacher Scholarship Loan Program (VTSPL)	2001 - Ongoing	Yes	No
Washington	Alternative Routes to Teaching Conditional Scholarship	2001 - Ongoing	Yes	No
Washington	Educator Retooling Conditional Scholarship	2007 - Ongoing	Yes	No
Washington	Teacher Shortage Conditional Grant Program	2016 - Ongoing	Yes	No
West Virginia		2010 - Ongoing	Barely	No

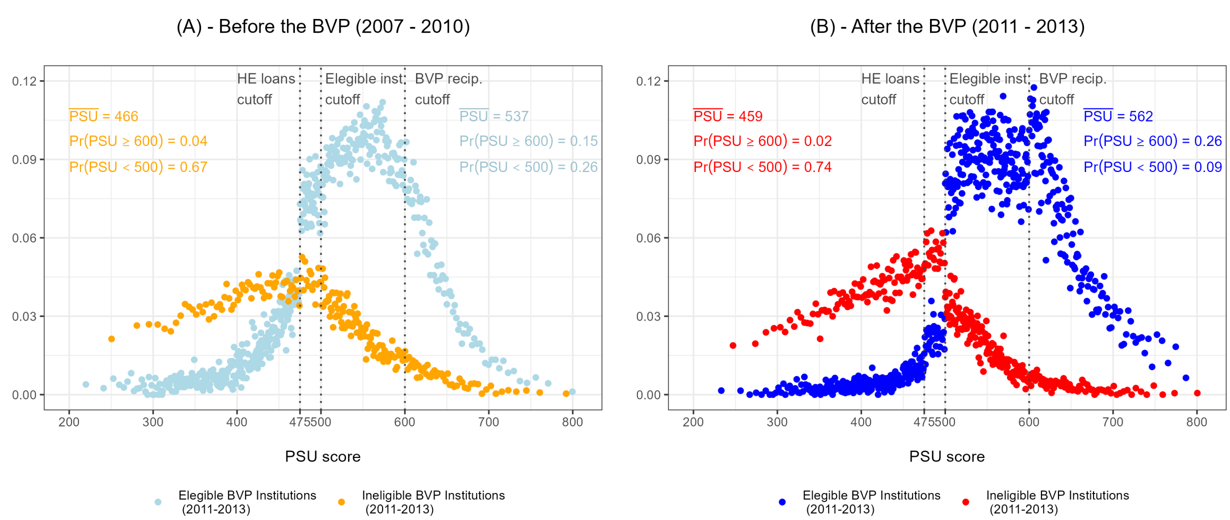
Notes. * Could be previous to this date.

A.2 Appendix Figures

Figure A1: BVP official Advertisement

Note: In the bottom : "Thousands of children want to learn from you. Enroll into a teaching program. With a score of 600 or more on the PSU, you can study for free. Teacher's Vocation Scholarship."

Figure A2: Teaching Program Enrollment Over Time



A.3 Appendix Tables

Table A4: Sample Mean Characteristics (Expanded)

	All Obs N=731,269 $\overline{PSU}$=503 (1)	All BVP N=7,742 $\overline{PSU}$=635 (2)	LLR. Mean at 600 points: 600 points cutoff (3)	Always TP.Stdnts (4)	Always Takers (5)	All Compliers (6)	Decomposition: Marginal TP.Stdnts Complier (7)	Inf.marginal TP.Stdnts Complier (8)
			100%	8.9%	.7%	9.2%[c]	3.1%	6%
				[.083;.095]	[.006;.009]	[.087;.096]	[.023;.039]	[.053;.067]
Female	.527	.544	.494	.588	.729	.557	.448	.614
	[.526;.528]	[.533;.555]	[.491;.499]	[.566;.606]	[.64;.822]	[.535;.59]	[.274;.59]	[.573;.716]
Age	18.91	19.73	18.75	19.29	17.86	19.86	21.12	19.2
	[18.9;18.92]	[19.65;19.82]	[18.75;18.78]	[19.16;19.44]	[17.69;18.02]	[19.69;20.15]	[19.85;22.42]	[18.8;19.72]
Private Health Insurance	.215	.292	.358	.252	.115	.248	.216	.264
	[.214;.216]	[.282;.302]	[.352;.361]	[.237;.276]	[.068;.162]	[.217;.265]	[.086;.302]	[.221;.299]
Lower Income (≈ <Q50)	.515	.389	.338	.426	.505	.436	.516	.394
	[.513;.516]	[.378;.4]	[.332;.34]	[.399;.446]	[.418;.591]	[.414;.471]	[.398;.672]	[.324;.446]
Paid Job on Test Year	.086	.118	.065	.099	.027	.12	.134	.112
	[.085;.086]	[.111;.125]	[.062;.066]	[.084;.113]	[0;.05]	[.104;.139]	[.038;.213]	[.084;.155]
Father Education	11.61	12.91	13.17	12.42	11.4	12.56	12.83	12.42
	[11.6;11.62]	[12.83;12.99]	[13.17;13.22]	[12.26;12.6]	[10.69;12.11]	[12.31;12.73]	[11.81;13.51]	[12.09;12.84]
Mother Education	11.47	12.63	12.92	12.25	11.01	12.24	12.35	12.18
	[11.47;11.48]	[12.55;12.7]	[12.91;12.96]	[12.11;12.4]	[10.22;11.81]	[11.98;12.38]	[11.41;13.25]	[11.7;12.44]
Capital City	.375	.485	.429	.404	.381	.456	.53	.418
	[.374;.377]	[.474;.497]	[.422;.431]	[.382;.434]	[.301;.461]	[.423;.475]	[.364;.642]	[.367;.473]
Public High School	.328	.289	.226	.273	.395	.275	.283	.27
	[.327;.329]	[.279;.299]	[.219;.226]	[.253;.293]	[.315;.475]	[.247;.293]	[.154;.376]	[.237;.322]
Voucher Private School	.564	.549	.577	.621	.605	.615	.596	.62
	[.563;.565]	[.538;.56]	[.577;.589]	[.6;.637]	[.514;.697]	[.597;.654]	[.495;.75]	[.57;.68]
Private Unsubsidized School	.108	.162	.192	.111	.019	.109	.133	.1
	[.107;.109]	[.154;.171]	[.184;.194]	[.099;.127]	[-.002;.039]	[.087;.123]	[.042;.198]	[.07;.12]
High School GPA Score	539	606	590	582	673	574	569	577
	[539;539]	[604;608]	[589;591]	[578;587]	[655;692]	[567;576]	[544;591]	[564;583]
N° of PSU Test Attempts	1.31	1.62	1.5	1.63	1.15	1.63	1.75	1.57
	[1.31;1.31]	[1.61;1.64]	[1.48;1.5]	[1.6;1.66]	[1.06;1.24]	[1.58;1.67]	[1.51;1.96]	[1.47;1.67]

Notes. (a): All the characteristics come from DEMRE PSU pre-test survey and DEMRE PSU data-sets. **(b):** Columns (3) to (8) are Linear Local Regressions (LLR) mean estimates at the 600 points cutoff. (2): Linear Local Regression (LLR) non-parametric mean based on Calonico et al. (2018) (3)-(8): LLR non-parametric mean based on Calonico et al. (2014) LLR method (3): Right-side of 600 cutoff of test-takers not enrolled at a Teaching Program (TP). (4): Left-side of the cutoff of test-takers enrolled at a TP. (5): BVP recipients at the left of the cutoff. (6) BVP Compliers mean characteristic. (7): BVP Compliers induced to enroll at a TP by the program at 600 mean characteristic. (8): BVP Compliers induced to take the program but who would enroll at a TP regardless of the program mean characteristic. (7) and (8) Are based on adapting Comey et al. (2024) to a Fuzzy RDD setting. **(c):** The share of marginal compliers equals 3.148695% and compliers 9.154834% **(d):** Income variables come from brackets $\{1,2\}$, $\{3,4\}$, $\{5,...,12\}$ on the DEMRE family income question. **(e):** [.,.]: are 95% confidence intervals (CI). In (1) Standard normal CI. In (2) non-parametric robust CI using Calonico et al. (2018) . In (3) - (7) non-parametric robust CI using Calonico et al. (2014) . In (8) robust CI using a wild-bootstrap procedure adapted from He and Bartalotti (2020).

Table A5: Differences Between Mean Characteristics

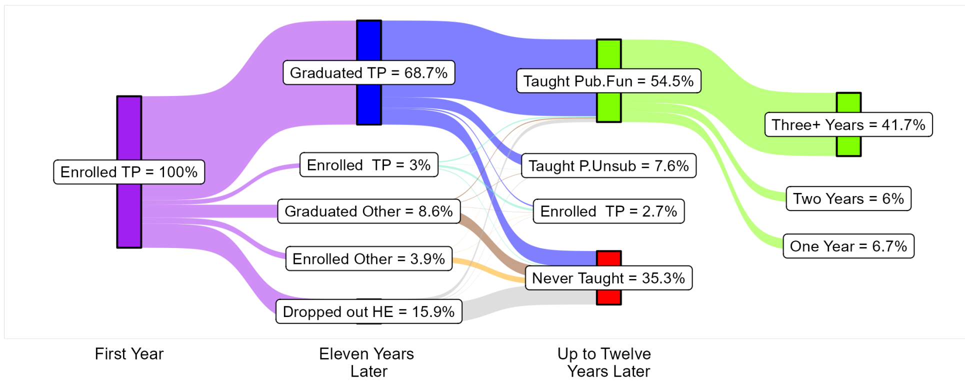
	Always TP. Students -					Always Takers -	Marg. TP Compliers -
	Never TP.Students (4) - (*)	Always Takers (4) - (5)	All Compliers (4) - (6)	Marg. TP Compliers (4) - (7)	Inframarg. TP Comp. (4) - (8)	All Compliers (5) - (6)	Inframarg. TP Comp. (7) - (8)
Female	.098***	-.142***	.03	.14	-.027	.172***	-.166**
	[.079;.121]	[-.228;-.058]	[-.015;.055]	[-.062;.338]	[-.128;.017]	[.098;.267]	[-.435;-.005]
Age	.64***	1.43***	-.58***	-1.84***	.08	-2***	1.92**
	[.52;.79]	[1.24;1.65]	[-.92;-.36]	[-2.73;-.41]	[-.4;.44]	[-2.33;-1.79]	[.14;3.41]
Private Health Insurance	-.121***	.137***	.005	.036	-.012	-.132***	-.048
	[-.134;-.096]	[.097;.196]	[-.02;.045]	[-.041;.182]	[-.034;.023]	[-.186;-.066]	[-.174;.092]
Lower Income (aprx <Q50)	.102***	-.078	-.009	-.089**	.033	.069	.122
	[.071;.121]	[-.16;.01]	[-.049;.02]	[-.252;0]	[-.023;.08]	[-.035;.144]	[-.041;.365]
Paid Job on Test Year	.042***	.072***	-.021**	-.035	-.014	-.093***	.021
	[.026;.056]	[.05;.097]	[-.045;-.001]	[-.114;.053]	[-.052;.012]	[-.118;-.069]	[-.09;.139]
Father Education	-.87***	1.02***	-.15	-.41	-.01	-1.16***	.4
	[-1.07;-.7]	[.47;1.75]	[-.35;.13]	[-1.19;.65]	[-.33;.29]	[-2;-.48]	[-1;1.24]
Mother Education	-.75***	1.24***	.02	-.09	.07	-1.22***	.17
	[-.9;-.56]	[.5;1.9]	[-.18;.35]	[-.94;.75]	[-.14;.49]	[-2.02;-.54]	[-.85;1.48]
Capital City	-.022*	.023	-.052**	-.126	-.014	-.075	.113
	[-.043;.003]	[-.045;.117]	[-.072;-.001]	[-.242;.04]	[-.056;.03]	[-.165;.019]	[-.088;.257]
Public High School	.055***	-.122**	-.001	-.007	.003	.122***	.01
	[.032;.077]	[-.208;-.038]	[-.03;.039]	[-.105;.115]	[-.045;.031]	[.04;.224]	[-.169;.118]
Voucher Private School	.039***	.013	.007	.028	0	-.006	-.03
	[.018;.06]	[-.081;.083]	[-.042;.036]	[-.119;.154]	[-.05;.04]	[-.091;.09]	[-.19;.12]
Private Unsubsidized School	-.088***	.091***	.002	-.022	.01	-.089***	.04
	[-.104;-.072]	[.076;.11]	[-.015;.031]	[-.09;.066]	[0;.04]	[-.12;-.063]	[-.05;.12]
High School GPA Score	-9***	-91***	8***	13	6**	99***	-7
	[-14;-5]	[-106;-75]	[3;17]	[-11;39]	[0;18]	[87;121]	[-35;25]
N of PSU Test Attempts	.15***	.48***	0	-.12	.06	-.48***	.18
	[.11;.18]	[.41;.57]	[-.05;.06]	[-.34;.11]	[-.03;.15]	[-.57;-.38]	[-.12;.45]

Notes. (a) (4) - (8) numbering corresponds to table 1 estimates (b) [.,.] : 95% bootstrap mean confidence intervals considering non-parametric mean adapted from He and Bartalotti (2020) . To the right of the second column are also robust 95% confidence intervals estimated using a Linear Local Regression, a uniform kernel, and a bandwidth based on auxiliary variables and regression discontinuity of the non-parametric designs from the previous table and the methods of Calonico et al. (2014) . b) (*) numbering corresponds to Never Teaching Program (TP) Students. These are all high-school graduates at 600 points who would have never studied a teaching program in the area. It corresponds to close to 88% of the sample.

B Online Appendix

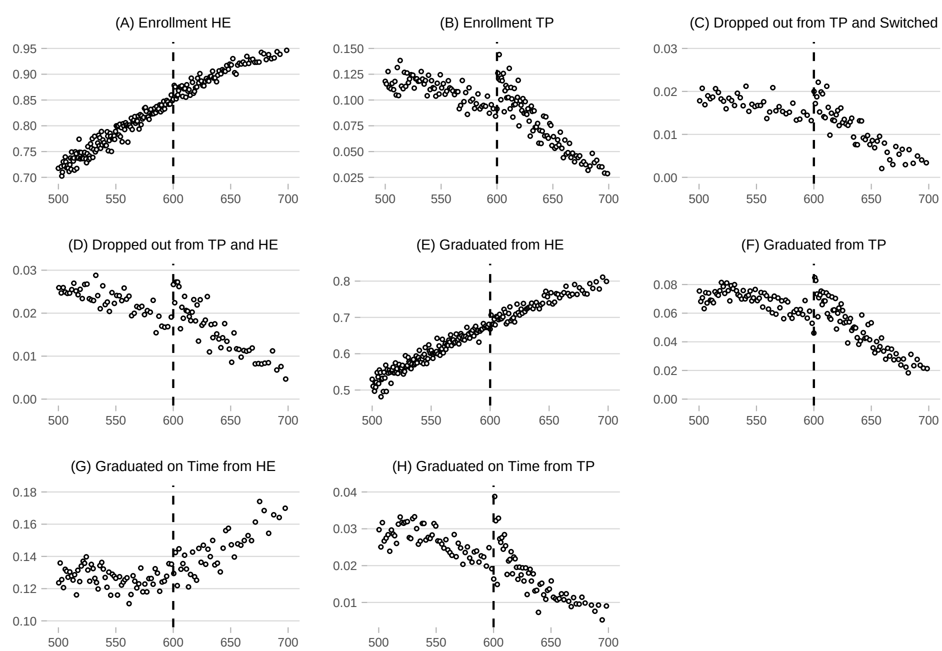
B.1 Online Appendix Figures

Figure B1: High-Scoring Students' Teaching Program Trajectory
Seven Years Before the Policy Implementation (2004 cohort)



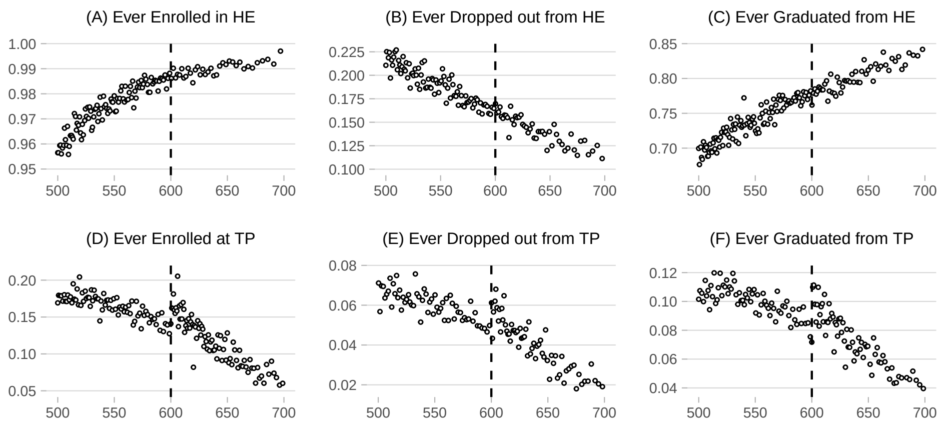
Notes. (1) High-Scoring Students: Students scoring $\geq$ 600 at admissions, HE: Higher Education, Other: Programs $\neq$ from Teaching, T.hrs: Teaching hours per year P.Unsub: Private unsubsidized schools (2) Outcomes are prioritized from top to bottom. Eg. Graduation from a teaching program counts as "Graduated" no matter of whether the student enrolled in another teaching program, enrolled or graduated in non non-teaching program as well.

Figure B2: Probabilities of Immediately Enrolling and Dropping Out or Graduating Following Immediate Enrollment after Taking the PSU



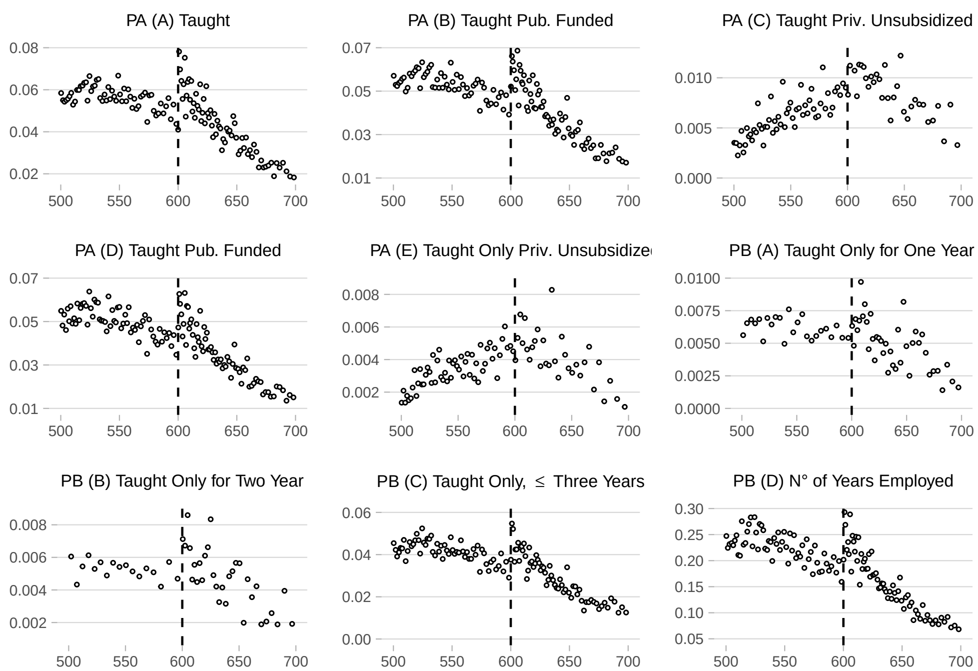
Notes. The non-parametric means of test-takers higher-education outcomes are built using IMSE-optimal quantile-spaced method using polynomial regression implemented in the rdplot R command described by Calonico et al. (2015a) and Calonico et al. (2015b).

Figure B3: Probabilities of Ever Enrolling, Dropping out or Graduating after Taking the PSU



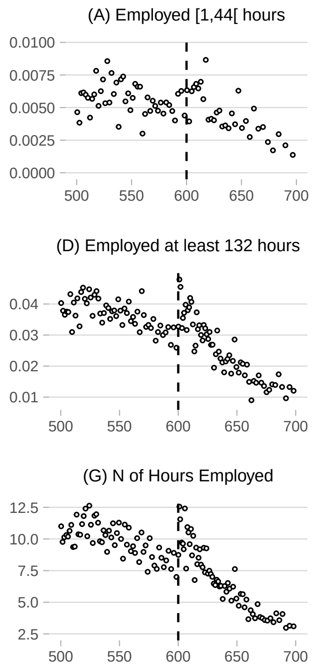
Notes. Non-parametric means of test-takers higher-education outcomes are built using IMSE-optimal quantile-spaced method using polynomial regression implemented in the rdplot R command described by Calonico et al. (2015a) and Calonico et al. (2015b).

Figure B4: Probability of Teaching by Type of School (PA) and by Year in in Pub. Fund. School (PB)



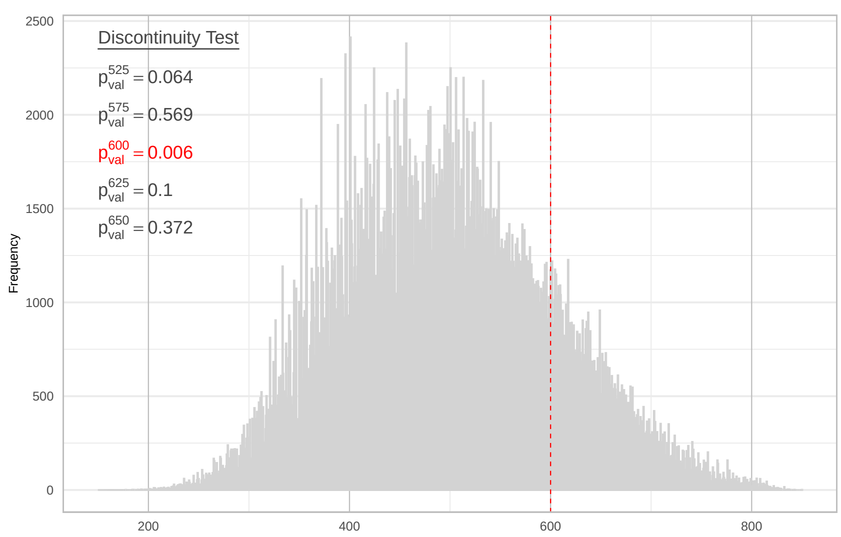
Notes. All graphs illustrate teaching Outcomes after Immediate Enrollment in Teaching Program after the PSU. The non-parametric means of test-takers higher-education outcomes are built using IMSE-optimal quantile-spaced method using polynomial regression implemented in the rdplot R command described by Calonico et al. (2015a) and Calonico et al. (2015b).

Figure B5: Conditional Mean of Teaching in Pub. Fund. School by Contract Hours (PC)



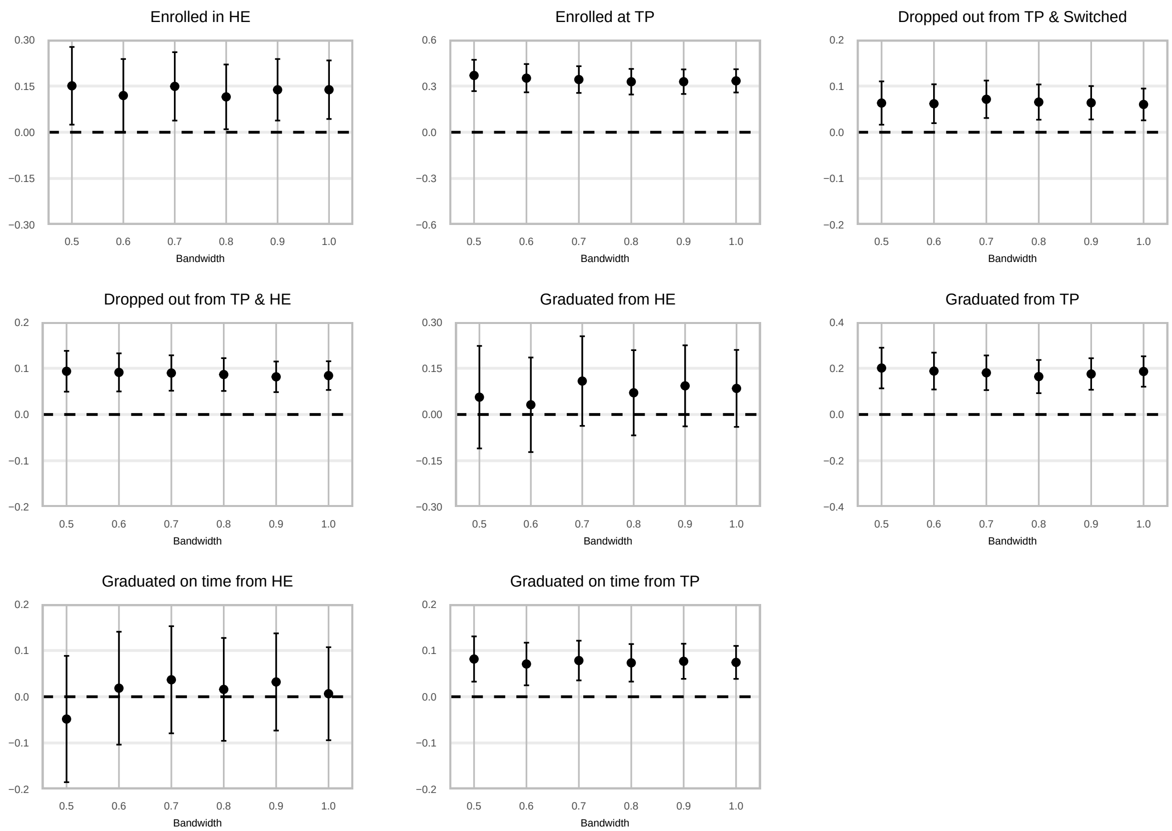
Notes. All graphs illustrate teaching Outcomes after Immediate Enrollment in Teaching Program after the PSU. The non-parametric means of test-takers higher-education outcomes are built using IMSE-optimal quantile-spaced method using polynomial regression implemented in the rdplot R command described by Calonico et al. (2015a) and Calonico et al. (2015b).

Figure B6: PSU Score Histogram and Manipulation Test Results



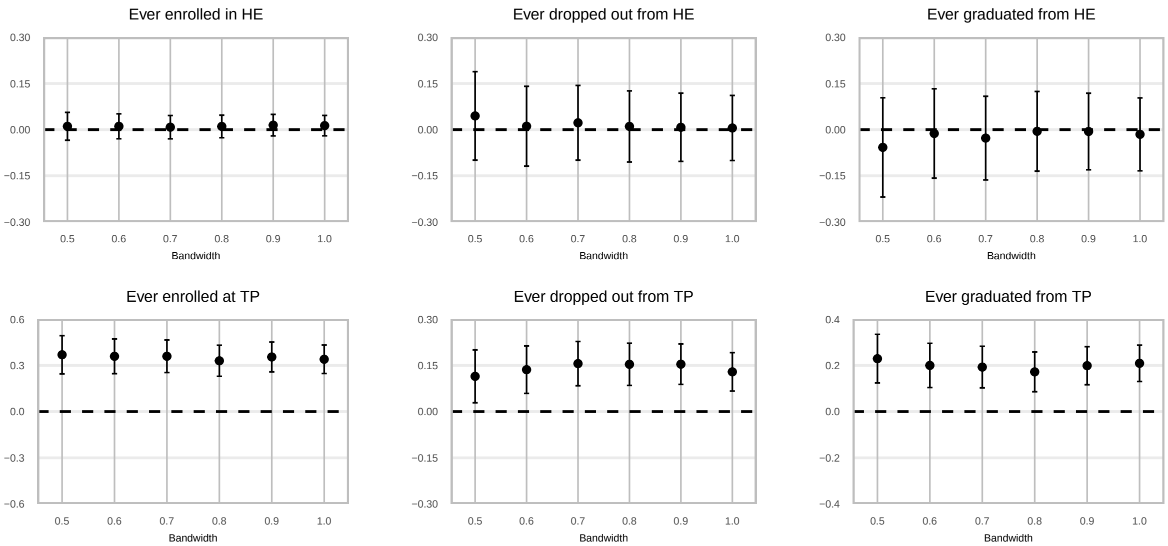
Notes. (i) The main goal of this Figure is to illustrate that density tests at 600 points, and other nearby cutoffs unrelated to other policies, fail to reject the hypothesis of a continuous density distribution. Because this discontinuity is generally interpreted as potential manipulation, this article also considers donut hole robust checks, although it is highly unlikely that manipulation may be a problem. The tests is administered by an independent institution, and students are regrouped far away from their school peers without knowing how much each score values. (ii) The density discontinuity test P values are estimated using the rddensity package discussed in Cattaneo et al. (2018) for the program requirement score in red, and placebo cutoffs. (iii) Each bin is 0.5 PSU points, the smallest bin possible which amounts to 0.005 Standard Deviations.

Figure B7: Test-Takers Higher Education Effects - Panel A - Different Bandwiths



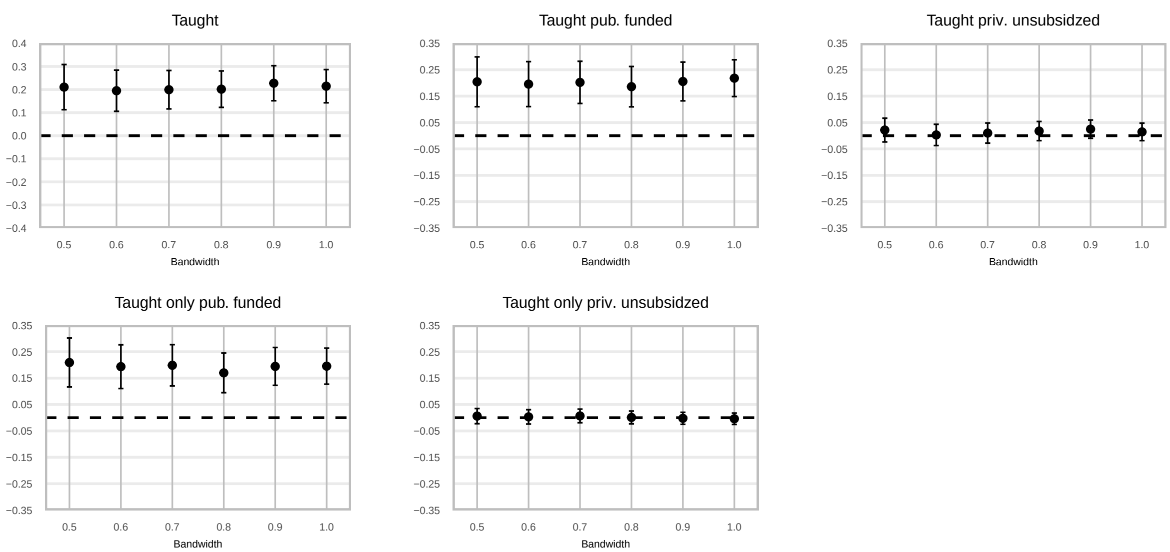
Notes. These figures illustrate the effects of the BVP on test-takers using fuzzy RDD. The x-axis correspond to different bandwidths chosen as multiples of the MSE optimal bandwidths computed following Calonico et al. (2014). The points illustrate the estimated effects, and the lines depict the 95% confidence intervals.

Figure B8: Test-Takers Higher Education Effects - Panel B - Different Bandwiths



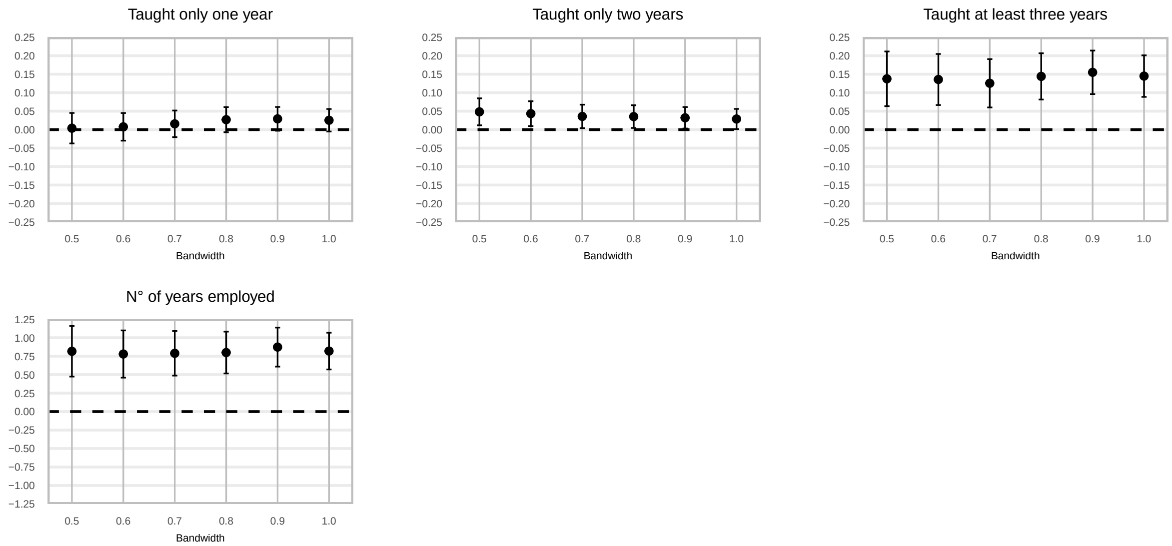
Notes. These figures illustrate the effects of the BVP on test-takers using fuzzy RDD. The x-axis corresponds to different bandwidths chosen as multiples of the MSE optimal bandwidths computed following Calonico et al. (2014). The points illustrate the estimated effects, and the lines depict the 95% confidence intervals.

Figure B9: Teaching Effects - Panel A - Different Bandwiths



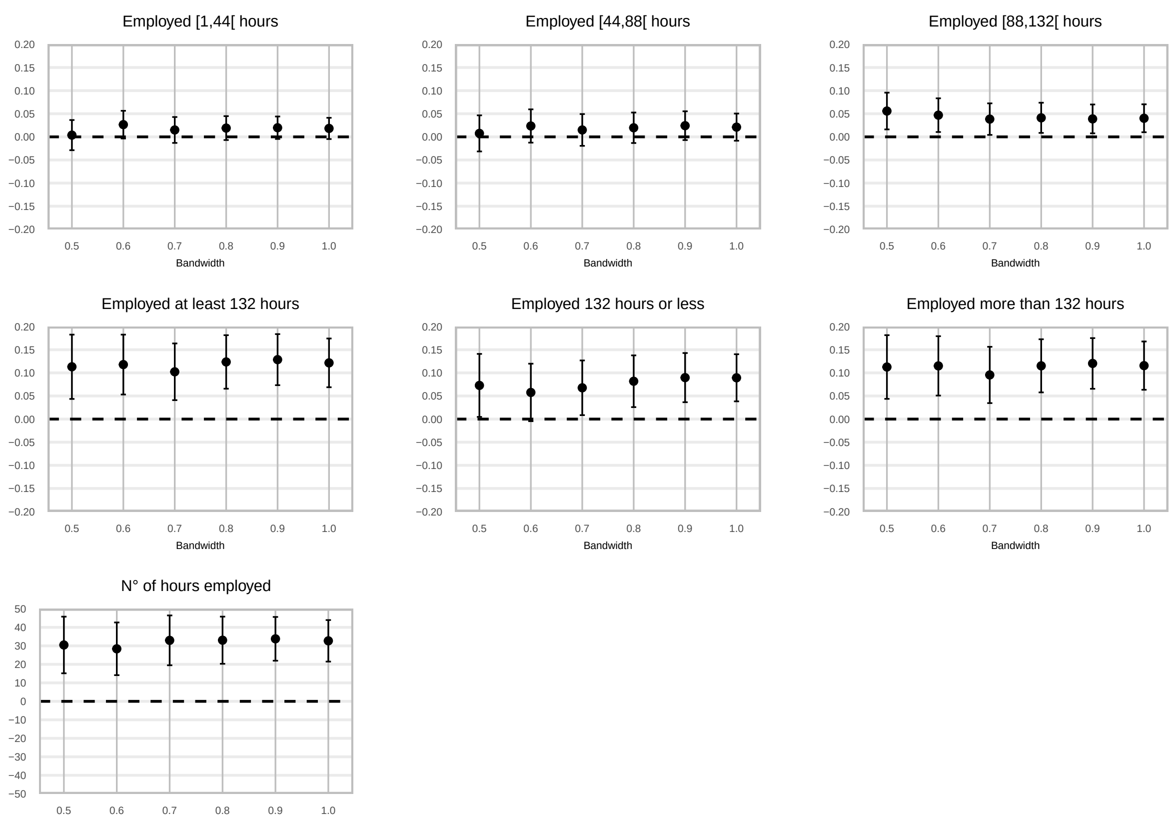
Notes. These figures illustrate the effects of the BVP on test-takers using fuzzy RDD. The x-axis corresponds to different bandwidths chosen as multiples of the MSE optimal bandwidths computed following Calonico et al. (2014). The points illustrate the estimated effects, and the lines depict the 95% confidence intervals.

Figure B10: Teaching Effects - Panel B - Different Bandwiths



Notes. These figures illustrate the effects of the BVP on test-takers using fuzzy RDD. The x-axis corresponds to different bandwidths chosen as multiples of the MSE optimal bandwidths computed following Calonico et al. (2014). The points illustrate the estimated effects, and the lines depict the 95% confidence intervals.

Figure B11: Teaching Effects - Panel C - Different Bandwiths



Notes. These figures illustrate the effects of the BVP on test-takers using fuzzy RDD. The x-axis corresponds to different bandwidths chosen as multiples of the MSE optimal bandwidths computed following Calonico et al. (2014). The points illustrate the estimated effects, and the lines depict the 95% confidence intervals.

B.2 Online Appendix Tables

Table B1: Balance Tests

	Mean Below 600 Baseline	Est.	Std. Dev	p val	Robust Inference: Conf. Int	Robust Inference: Band-width	Eff. Number Obs.
Pre-Higher Education							
Female	.491	.008	(.005)	.142	[-.003;.021]	73.8	147,395
Age	18.76	.02	(.025)	.474	[-.036;.078]	83.1	171,490
Private Health Insurance	.354	.003	(.005)	.881	[-.011;.013]	73.6	124,348
Lower Income (aprx <Q50)	.333	.005	(.005)	.159	[-.003;.019]	95.9	126,221
Paid Job on Test Year	.065	-.001	(.002)	.980	[-.005;.005]	104	171,490
Father Education	13.2	-.02	(.041)	.401	[-.130;.052]	70.5	104,772
Mother Education	12.98	-.09***	(.039)	.009	[-.192;-.028]	72.3	103,002
Capital City	.42	.004	(.005)	.682	[-.009;.013]	86.4	161,201
Public High School	.22	.001	(.004)	.894	[-.008;.010]	91.4	180,115
Voucher Private School	.581	.003	(.006)	.362	[-.007;.018]	72.4	113,519
Private Unsubsidized School	.188	.001	(.004)	.943	[-.011;.010]	63.7	122,464
High School GPA Score	589	.236	(.933)	.992	[-2.10;2.10]	65.1	127,460
N° of PSU Test Attempts	1.49	.01	(.007)	.329	[-.008;.024]	102.1	181,811
PSU Scores ≠ Math or Language							
Takes Hist. and Soc.	.52	.01*	(.005)	.097	[0;.020]	66.7	136,937
Hist. and Soc. score	606	-.17	(1.041)	.825	[-2.70;2.10]	59.5	64,847
Takes Sci.	.703	-.005	(.005)	.587	[-.015;.009]	62.1	120,716
Sci. score	574	-1.24*	(.74)	.051	[-3.14;.010]	63.3	67,997
Year							
Year 2011	.333	-.002	(.005)	.976	[-.011;.012]	71	131,263
Year 2012	.33	-.004	(.004)	.275	[-.010;0]	98.6	180,115
Year 2013	.335	.002	(.005)	.696	[-.009;.014]	67.9	139,941

Notes. (1) Mean Below 600: The mean below the cutoff of the local linear ITT without controls and with a uniform kernel. Est.: ITT estimate of the effect of crossing the 600 point threshold on pre-higher education characteristics. Std.Dev: Standard Deviation of the Estimate. p val: P-value of the estimate. Robust Inference: Non-parametric robust inference following Calonico et al. (2014). Conf. Int: Confidence Interval at a significance level of 5 percent. Bandwidth: MSE Bandwith following the procedure suggested by Calonico et al. (2014). Eff Number Obs: Number of observations within the bandwidth. (2) PSU scores standard deviations are always close to 100 points by design. Mother and Father Education are years of higher education where 12 years is the number of years required to finish high-school.

Table B2: Test-Takers Higher Education Effects (Robustness Checks)

	Baseline Specification (1)	Quadratic Polynomial (2)	Triangular Kernel (3)	No controls (4)	Donut Hole (5)	Donut Hole2 (6)
Panel A - Outcomes requiring enrollment the year after taking the PSU						
Enrolled in HE	0.114**	0.129**	0.139	0.127*	0.143	0.143
	(0.053)	(0.054)	(0.059)	(0.054)	(0.102)	(0.102)
Enrolled at TP	0.328***	0.340***	0.339***	0.346***	0.317**	0.317***
	(0.039)	(0.04)	(0.043)	(0.04)	(0.075)	(0.075)
Dropped out from TP & Switched	0.063**	0.069**	0.067**	0.069**	0.057	0.057
	(0.019)	(0.019)	(0.021)	(0.019)	(0.036)	(0.036)
Dropped out from TP & HE	0.091***	0.082***	0.088**	0.103**	0.084	0.084**
	(0.021)	(0.022)	(0.024)	(0.022)	(0.039)	(0.039)
Graduated from HE	0.073	0.101	0.060	0.086	0.286	0.286**
	(0.07)	(0.072)	(0.078)	(0.072)	(0.135)	(0.135)
Graduated from TP	0.174***	0.192***	0.182***	0.174***	0.175*	0.175***
	(0.035)	(0.035)	(0.038)	(0.035)	(0.067)	(0.067)
Graduated on time from HE	0.006	0.008	0.006	0.029	0.189	0.189*
	(0.051)	(0.053)	(0.057)	(0.053)	(0.098)	(0.098)
Graduated on time from TP	0.076**	0.079**	0.075**	0.079**	0.111*	0.111***
	(0.022)	(0.023)	(0.025)	(0.023)	(0.043)	(0.043)
Panel B - Outcomes requiring enrollment at any year after taking the PSU						
Ever Enrolled in HE	0.015	0.020	0.009	0.015	0.004	0.004
	(0.017)	(0.018)	(0.019)	(0.018)	(0.034)	(0.034)
Ever Dropped Out from HE	-0.003	0.001	0.017	0.005	-0.084	-0.084
	(0.055)	(0.057)	(0.062)	(0.057)	(0.106)	(0.106)
Ever Graduated from HE	0.002	0.039	-0.022	0.000	0.129	0.129
	(0.062)	(0.063)	(0.069)	(0.065)	(0.119)	(0.119)
Ever Enrolled at TP	0.344***	0.356***	0.346***	0.365***	0.325	0.325***
	(0.048)	(0.049)	(0.053)	(0.049)	(0.093)	(0.093)
Ever Dropped Out from TP	0.138***	0.142***	0.138*	0.157**	0.148	0.148**
	(0.033)	(0.034)	(0.037)	(0.033)	(0.062)	(0.062)
Ever Graduated from TP	0.204***	0.217***	0.192***	0.205***	0.218	0.218***
	(0.041)	(0.042)	(0.045)	(0.042)	(0.079)	(0.079)

Notes. (i) HE: Higher Education, TP: Teaching Program. (ii) (1) Baseline Specification: The same fuzzy specification from the main table using a linear local regression and a uniform kernel. (2) Quadratic Polynomial: Polynomial fuzzy specification with a uniform kernel. The bandwidth is built in the same way as for the baseline specification, but with a quadratic polynomial specification. That is, the bandwidth is the minimum bandwith across all the outcomes with a quadratic polynomial. (3) Triangular kernel: Linear local fuzzy specification and a triangular kernel. The bandwith is built in the same way as for the baseline specification, but with a triangular kernel. (4) Baseline specification but without controls. The bandwith is built in the same way as for other columns. (iii) Heteroscedasticity robust standard-errors in parenthesis. (5) Equals to the first column but dropping all the units around 600 points within one quarter of the bandwith (6) It's the same as (5) but p-values do not consider non-parametric bias (iv) *p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01 of estimates considering non-parametric bias and heteroscedasticity using the robust estimates of Calonico et al. (2014).

Table B3: Teaching Effects (Robustness Checks)

	Baseline Specification (1)	Quadratic Polynomial (2)	Triangular Kernel (3)	No controls (4)	Donut Hole (5)	Donut Hole2 (6)
Panel A - Taught by Type of school						
Taught	0.181***	0.199***	0.186***	0.189***	0.196**	0.196***
	(0.033)	(0.032)	(0.031)	(0.033)	(0.061)	(0.061)
Taught pub. funded	0.178***	0.198***	0.181***	0.182***	0.196**	0.196***
	(0.032)	(0.031)	(0.03)	(0.031)	(0.058)	(0.058)
Taught priv. unsubsidzed	0.021	0.023	0.019	0.026	0.025	0.025
	(0.015)	(0.015)	(0.015)	(0.015)	(0.028)	(0.028)
Taught only pub. funded	0.160***	0.176***	0.167***	0.163***	0.172**	0.172***
	(0.031)	(0.03)	(0.029)	(0.03)	(0.057)	(0.057)
Taught only priv. unsubsidzed	0.003	0.001	0.005	0.007	0	0
	(0.011)	(0.011)	(0.01)	(0.011)	(0.021)	(0.021)
Panel B - Taught in publicly funded schools: years						
Taught only one year	0.021	0.018	0.016	0.019	0.055	0.055**
	(0.012)	(0.012)	(0.011)	(0.012)	(0.023)	(0.023)
Taught only two years	0.021**	0.027	0.023**	0.023**	0.008**	0.008
	(0.012)	(0.012)	(0.011)	(0.012)	(0.021)	(0.021)
Taught at least three years	0.136***	0.153***	0.143***	0.140***	0.133	0.133**
	(0.029)	(0.028)	(0.027)	(0.028)	(0.053)	(0.053)
N° of years employed	0.719***	0.794***	0.740***	0.731***	0.798**	0.798***
	(0.146)	(0.142)	(0.138)	(0.143)	(0.27)	(0.27)
Panel C - Taught in publicly funded schools: contract hours						
Employed [1,44[hours	0.008	0.008	0.006	0.008	0.028	0.028
	(0.012)	(0.011)	(0.011)	(0.011)	(0.022)	(0.022)
Employed [44,88[hours	0.018*	0.020	0.019	0.018	-0.002	-0.002
	(0.012)	(0.012)	(0.012)	(0.012)	(0.022)	(0.022)
Employed [88,132[hours	0.019**	0.022*	0.023**	0.018**	0.009	0.009
	(0.013)	(0.012)	(0.012)	(0.012)	(0.025)	(0.025)
Employed at least 132 hours	0.125***	0.143***	0.126***	0.130***	0.149	0.149***
	(0.027)	(0.026)	(0.025)	(0.026)	(0.049)	(0.049)
Employed 132 hours or less	0.050**	0.055*	0.053**	0.050**	0.047	0.047
	(0.021)	(0.021)	(0.02)	(0.02)	(0.039)	(0.039)
Employed more than 132 hours	0.119***	0.137***	0.122***	0.124***	0.137*	0.137***
	(0.027)	(0.026)	(0.025)	(0.026)	(0.049)	(0.049)
N° of hours employed	29.8***	33.3***	30.6***	30.5***	35.4*	35.4***
	(6.467)	(6.314)	(6.103)	(6.355)	(11.906)	(11.906)

Notes. (i) HE: Higher Education, TP: Teaching Program. (ii) (1) Baseline Specification: The same fuzzy specification from the main table using a linear local regression and a uniform kernel. (2) Quadratic Polynomial: Polynomial fuzzy specification with a uniform kernel. The bandwith is built in the same way as for the baseline specification but with a quadratic polynomial specification. That is, the bandwith is the minimum bandwith across all the outcomes with a quadratic polynomial. (3) Triangular kernel: Linear local fuzzy specification and a triangular kernel. The bandwith is built in the same way as for the baseline specification but with a triangular kernel. (4) Baseline specification but without controls. The bandwith is built in the same way as for other columns. (iii) Heteroscedasticity robust standard errors in parenthesis. (5) Equals to the first column but dropping all the units around 600 points within one quarter of the bandwith (6) It's the same as (5) but p-values do not consider non-parametric bias (iv) *p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01 of estimates considering non-parametric bias and heteroscedasticity using the robust estimates of Calonico et al. (2014).

Table B4: Teaching Effects Following Immediate Enrollment vs Ever Enrollment

	Immediate Enrollment		Ever Enrollment	
	Potential No Treatment Mean Baseline (1)	Fuzzy RDD Treatment Estimate (2)	Potential No Treatment Mean Baseline (3)	Fuzzy RDD Treatment Estimate (4)
Panel A - Taught by Type of School				
Taught	0.375***	0.181***	0.381***	0.197***
	(0.028)	(0.033)	(0.033)	(0.038)
Taught pub. funded	0.339***	0.178***	0.345***	0.194***
	(0.026)	(0.032)	(0.031)	(0.036)
Taught priv. unsubsidzed	0.067***	0.021	0.073***	0.017
	(0.012)	(0.015)	(0.014)	(0.017)
Taught only pub. funded	0.308***	0.160***	0.308***	0.180***
	(0.025)	(0.031)	(0.029)	(0.035)
Taught only priv. unsubsidzed	0.036***	0.003	0.036	0.003
	(0.009)	(0.011)	(0.011)	(0.013)
Panel B - Taught in Publicly Funded Schools: Years				
Taught only one year	0.039***	0.021	0.035**	0.023
	(0.009)	(0.012)	(0.014)	(0.017)
Taught only two years	0.041**	0.021**	0.044**	0.037**
	(0.008)	(0.012)	(0.012)	(0.015)
Taught at least three years	0.258***	0.136***	0.266***	0.134***
	(0.022)	(0.029)	(0.025)	(0.03)
N° of years employed	1.386***	0.719***	1.41***	0.746***
	(0.118)	(0.146)	(0.127)	(0.152)
Panel C - Taught in Publicly Funded Schools: Contract Hours				
Employed [1,44[hours	0.040***	0.008	0.027*	0.015
	(0.009)	(0.012)	(0.013)	(0.015)
Employed [44,88[hours	0.052***	0.018*	0.060***	0.016
	(0.009)	(0.012)	(0.013)	(0.016)
Employed [88,132[hours	0.045**	0.019**	0.036*	0.040**
	(0.009)	(0.013)	(0.012)	(0.015)
Employed at least 132 hours	0.223***	0.125***	0.237***	0.115**
	(0.021)	(0.027)	(0.023)	(0.028)
Employed 132 hours or less	0.138***	0.050**	0.123***	0.078
	(0.016)	(0.021)	(0.022)	(0.027)
Employed more than 132 hours	0.221***	0.119***	0.237***	0.108**
	(0.021)	(0.027)	(0.022)	(0.028)
N° of hours employed	61.487***	29.753***	62.998***	30.299**
	(05.212)	(06.467)	(05.616)	(06.728)

Notes. This tables illustrates the strong similarity between estimates of the teaching effect considering immediate enrollment against ever-enrolled estimates. Immediate enrollment estimates are the interaction between any outcome with a binary variable that indicates enrollment in higher education right after taking the test combined with a treatment outcome defined as taking the BVP right after the test for immediate enrollment or ever taking the BVP for ever enrollment. (i) HE: Higher Education, TP: Teaching Program. (ii) (1) Potential No Treatment Mean Baseline: The average potential outcome of compliers at the cutoff (600 points) if they had not been treated. (2) Fuzzy estimates (3) ATPs Mean (Always Teaching Program students): The mean of students enrolled at a TP below the cutoff, estimated as the mean below the cutoff of a sharp RDD without controls in a sample restricted to teaching program students. (4) Immediate TP Benchmark: This is equal to 0.330, the fuzzy estimate of the effect on the proabability of enrolling at a TP, times column (3) for teaching program outcomes following immediate enrollment. The benchmark shows how would be the fuzzy estimate if the policy only attracted more students similar to teaching program students who immediately enrolled after scoring slightly below 600 points (iii) All the RDD estimates use a local linear regression, a uniform kernel and a 21.9 bandwidth, the minimum bandwith that minimizes the MSE following Calonico et al. (2014) across all the outcomes. (iv) Heteroscedasticity robust standard errors in parenthesis. (v) *p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01 of estimates considering non-parametric bias and heteroscedasticity using the robust estimates of Calonico et al. (2014).

Table B5: Effects on Ever Taking Competing Grants and Loans (MSE Bandwiths)

	Mean Below 600 Baseline (1)	ITT RDD Estimates (2)	Potential No Treatment Mean Baseline (3)	Fuzzy RDD Treatment Estimates (4)
Panel A - Grants				
Juan Gomez Millas	0.147***	-0.001	0.049	-0.011
	(0.006)	(0.005)	(0.055)	(0.054)
Bicentenario	0.366***	-0.053***	0.604***	-0.566***
	(0.008)	(0.006)	(0.07)	(0.062)
Other grants	0.328***	-0.022***	0.422***	-0.242***
	(0.006)	(0.005)	(0.059)	(0.056)
Any Grant	0.585***	-0.045***	0.681***	-0.469***
	(0.008)	(0.006)	(0.072)	(0.06)
Panel B - Loans				
Traditional University Loan	0.280***	-0.017***	0.397***	-0.176***
	(0.005)	(0.004)	(0.048)	(0.046)
State Guaranteed Loan	0.420***	-0.007	0.286*	-0.076
	(0.007)	(0.006)	(0.064)	(0.064)
Any Loan	0.610***	-0.024**	0.612***	-0.258***
	(0.006)	(0.006)	(0.06)	(0.059)

Notes. (i) These effects are for competing grants and loans in any year after taking the test, but unlike the table in the main text, it uses a the optimal fuzzy bandwitdh for each rows. (ii) HE: Higher Education, TP: Teaching Program. (ii) (1) Mean Below 600: The mean below the cutoff of the local linear ITT without controls. (2) ITT Estimates: Intent to Treat Estimates, using a sharp RDD with controls. (3) Potential No Treatment Mean Baseline: The average potential outcome of compliers at the cutoff (600 points) if they had not been treated. (4) Fuzzy estimates (iii) All the RDD estimates use a local linear regression, a uniform kernel and a bandwith that minimizes the MSE following Calonico et al. (2014) across all the outcomes. Each row uses the optimal bandwidth on the Fuzzy RDD estimates (iv) Heteroscedasticity robust standard-errors in parenthesis. (v) *p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01 of estimates considering non-parametric bias and heteroscedasticity using the robust estimates of Calonico et al. (2014).

Table B6: Annual Tuition Costs and Durations of Majors (2011-2015)

	BVP Eligible Teaching Majors			Non-Eligible Teaching Majors			Other Teaching Majors		
	p5	Mean	p95	p5	Mean	p95	p5	Mean	p95
Tuition costs in 2015 USD									
All students	2,328	3,225	5,370	1,035	2,139	3,799	1,372	3,332	6,576
PSU score near 600	2,328	3,163	4,953	1,795	3,106	4,692	2,312	4,520	6,739
Major duration in years									
All students	4	4.8	5.5	1.5	3.4	5	2	4	6
PSU score near 600	4	4.8	5	3.5	4.4	5	2	4.8	6

Notes. All statistics are weighted by the number of students in each major. p1 and 95 prefered to min/max because of suspicious outliers in the extremes. This does not affect mean results. PSU near 600 = +/- 50 pts. Source: Elaboration using SIES, DEMRE and MINEDUC data.

Table B7: Changes in Teaching Programs (TP)

	Before BVP (2008 - 2010)	After BVP (2011 - 2013)	Difference
Panel A - PSU average			
Elegible TP	537	562	25
Ineligible TP	466	459	-6
Any TP	511	522	11
Elegible TP - Ineligible TP	71	103	31
Panel B - Enrollment at TP Institution			
Elegible TP	64%	61%	-3%
Ineligible TP	36%	39%	3%
Any TP	100%	100%	0%
Panel C - Test-Takers Enrolled at a TP			
% Test-takers enroll. at a TP	7.9%	7.3%	-0.7%
$\bar{N}$ of Test-Takers enroll. at a TP	18,563	16,299	-2,264
$\bar{N}$ of Test-Takers	230,479	245,154	14,675

Notes. (1) PSU scores correspond to the Math and Language mean. (2) One Standard Deviation in the Math and Language mean is equal to 103 PSU points in 2008 - 2010 (3) Eligible institutions correspond to BVP eligible institutions at any moment between 2011 and 2013

B.3 Online Appendix Proofs

B.3.1 Proof of Counterfactual 2

This subsection explains how the $\beta_{TPcntr2}$ counterfactual is set up. The proof shows that the fuzzy treatment effect on outcomes Y—which require immediate enrollment in a teaching program and are equal to zero otherwise—is equal to the baseline times the ratio of the share of marginal compliers in teaching enrollment to the share of inframarginal teaching enrollment, under two assumptions.

Assumption B2.a - Outcome Monotonicity The first assumption is outcome monotonicity on teaching program enrollment which allows to decompose compliers C into compliers that are marginal in teaching program enrollment and those compliers who would have study teaching regardless of being slightly below or above the threshold of 600 points.

Assumption B2.b - Zero treatment effect effect on marginals and inframarginals The second assumption is that the treatement does not affect students who would have studied regardless, and the average outcome of Y on marginal students also equals the average outcome on Y on inframarginal students while untreated.

Under these two assumptions this subection proofs that $\beta_{TPcntr2}$:

$$E(Y(1)-Y(0)|C,S=600)=E(Y(0)|C,S=600)\frac{P(G=m|C,S=600)}{P(G=im|C,S=600)}$$

Proof

Note: All the expressions below are conditional on a score of $S = 600$, which is omitted for readability.

Because Y is restricted to outcomes that follow enrollment, marginal students -students who changed their higher education choice because of the treatment- have an an income equal to 0 on outcomes such as, graduation, dropping-out or teaching when untreated. Then given assumption Outcome Monotonicity assumption:

$$
\begin{aligned}
E(Y(0)|C) &= E(Y(0)|G = im, C)P(G = im|C) + 0 * P(G = m|C) \\
&= E(Y(0)|G = im, C)P(G = im|C)
\end{aligned}
$$

Which also means:

$$E(Y(0)|G = im, C) = \frac{E(Y(0)|C)}{P(G = im|C)} \tag{10}$$

Using the same logic:

$$E(Y(1) - Y(0)|C) = E(Y(1)|C) - E(Y(0)|G = im, C)P(G = im|C)$$

Given the zero treatment effect assumption on marginals and inframarginals:

$$E(Y(1) - Y(0)|C) = E(Y(0)|G = im, C) - E(Y(0)|G = im, C)P(G = im|C)$$

$$E(Y(1) - Y(0)|C) = E(Y(0)|G = im, C)P(G = m|C)$$

And finally, using Equation 10:

$$E(Y(1) - Y(0)|C) = \frac{E(Y(0)|C)}{P(G = im|C)}P(G = m|C)$$