Does Access to Student Loan Impact Higher Education Indicators?

Rebeca Regina Regatieri*

August 1, 2023

Abstract

The student loan policy is widely used in several countries and, in developing countries, is used to improve equity in access to higher education diplomas. Recent literature examines student loan impacts but faces limitations in robustness due to available databases, which don't include individuals interested in the program but not beneficiaries. The Student Financing Fund (Fies) is a Brazilian federal government student loan program that has existed for over two decades and, since 2015, has had a competitive selection process based on the score in the National High School Examination (Enem), which generates a passing grade for access to degrees and institutions offering vacancies in the program. I use a single database that crosses the administrative data identified in the records of the selection process and contract management of the Fies, with the data identified in the Enem and Census of Higher Education. Through the discontinuity regression methodology in the fuzzy design, I compared the academic evolution of individuals who applied in the program and had a performance slightly below the passing grade with individuals who applied and scored slightly above the passing grade. The Fies program significantly positively impacts higher education access (ranging from 51.2 p.p. to 31.5 p.p.) and interest in degree programs during the selection process (ranging from 41.1 p.p. to 25.2 p.p.). Additionally, it positively affects persistence in higher education, with a greater impact in the year following application (32 p.p.) and a lesser impact two to four years later (ranging from 21.8 p.p. to 13.7 p.p.). However, there is no effect on degree migration; Fies students do not change degrees more than the comparison group. The effects on degree completion are still being analyzed, and I await the results from the Inep secrecy room in August.due analysis.

Keywords: impact assessment, student loan, higher education

JEL Classification: C14, I23, I28.

*I thank Enlinson Mattos, my advisor in developing this work, who has been a great partner in the doctorate and this project. I also thank Fernanda Estevan, who is a co-supervisor in the development of this work. I am extremely grateful to the National Education Development Fund (FNDE), especially Rafael Rodrigues Tavares and Sylvia Cristina Toledo Gouveia, who supported me in formalizing and accessing Fies' administrative bases. I also thank Rafael for his frequent dialogues about the program.

1 Introduction

In the context of policies aimed at expanding higher education, the debate on the relevance of public student loan programs points out that low-income students, with qualification levels similar to high-income students, may not have access to higher education due to restrictions in the private credit market. Among other market failures, they do not have financial guarantees and cannot offer additional future income as collateral (Carneiro and Heckman, 2002; Carver et al., 2007; Brown et al., 2012; Sun and Yannelis, 2016; Lochner and Monge-Naranjo, 2016; Solis, 2017). The academic literature well documents the importance of state intervention to mitigate inequalities of opportunity in obtaining a higher education diploma (Becker, 1967; Sewell, 1971; Card, 1999; Katz et al., 1999; Belley et al., 2010; Bailey and Dynarski, 2011; Liu et al., 2016; Palmisano et al., 2022). In the early 2010s, fewer than 10% of people in the first quartile of Brazil's income distribution had higher education. In comparison, this percentage was over 70% in the last quartile.[1] In this sense, the objective of this work is to analyze whether students who obtained financing from the Student Financing Program (Fies), the public student credit program of the Brazilian federal government, were more likely to enroll in higher education and, in the degree of interest, to remain with active enrollment (retention), degree migration and to complete graduation in the ideal time.

In Brazil, part of the effort to expand access to higher education focused on instituting affirmative policies for admitting low-income students (graduates from public schools), Afro-descendant, and indigenous ethnic groups into public universities. From the mid-2000s onwards, policymakers implemented these policies more forcefully in the 2010s, especially with a 2012 federal law establishing quotas for access to higher education in federal public institutions. There was evidence of the positive impact of affirmative action policies in Brazil on access to higher education when the focus was on public school students or race (Francis and Tannuri-Pianto, 2012; Estevan et al., 2019; Vieira and Arends-Kuenning, 2019). The students admitted to public universities under affirmative action perform at similar levels to students who were not (Valente and Berry, 2017), and without much evidence of impact on the effort to prepare for selection processes (Estevan et al., 2019). There is evidence of impact on the choice of high schools (Mello, 2023), given that federal law established the quota for students from public schools.[2] However, the scope of the effects of inequality reduction policies in higher education through public institutions is limited because most students are in private institutions. In 2000, the percentage of enrollments in higher education in private institutions was around 67%, rising to around

[1] Data from the national household sample survey by the Brazilian Institute of Research Statistics.

[2] Within this quotas, there are also quotas for afro-descendants, indigenous people, and people with disabilities.

75% in 2017.[3]

Therefore, the other part of the Brazilian expansion effort focused on access to private higher education, with policies to subsidize tuition fees and grant student loans. There is evidence of the impact on access, retention, and completion in higher education of financial aid programs (non-refundable) to fund students, especially lower-income students (Dynarski, 2003; Fack and Grenet, 2015). On the refundable side, in more than the seventy countries where student loan is available, there are four types of objectives for these policies: i) equity, to increase access for students from a disadvantaged background; ii) financial, to raise funding for higher education institutions; iii) specific workforce training for the labor market, to overcome the shortage of professionals in certain careers or fields; and, iv) promotion of students' autonomy, giving them financial independence from their parents. The first objective is typical of developing countries, while the latter is common in developed countries (Dente and Piraino, 2011). In 2001, Brazil implemented Fies, a public student loan program run by the federal government, to reduce access inequality and contribute to the increase of graduates in some areas, such as teachers.

In the international empirical literature, there is a debate about how relevant the restriction to the credit market is in determining the inequality of access to higher education (Carneiro and Heckman, 2002; Belley and Lochner, 2007; Lochner and Monge-Naranjo, 2011): whether what matters to access higher education are short-term or long-term factors. The long-term factors are the cognitive development that occurs from the earliest years of age. Early childhood education and basic education define students' cognitive abilities in higher education admission processes. The prioritization of policies would be on long-term factors such as providing quality education and not on restrictions on the credit market (Cameron and Heckman, 2001; Carneiro and Heckman, 2002). As for the short-term factors, there are credit constraints to pay tuition and the opportunity cost of continuing to study instead of entering the job market. So, family income can determine the entry rate into higher education and the hours worked while studying (Kane, 1996; Belley and Lochner, 2007; Lochner and Monge-Naranjo, 2011; Solis, 2017).

More recent literature focusing on Chile's student loan program provides evidence of the impact on enrollment, retention, course completion and labor outcomes. Solis (2017) analyzes the impact of Chile's student loan programs on delays in college enrollment (college enrollment immediately after high school graduation), persistence (enrolled during the degree period), and the total number of college years, particularly for students with lower socioeconomic backgrounds. Montoya et al. (2018) use a regression discontinuity design to investigate if Chile's student loan program induces students to pursue college degrees that are more expensive and prolonged relative to technical education and the

[3]Statistics from the Higher Education Census, database better described later.

effects on labor outcome. Results found were that the program induced students to choose a university education that requires more years of training and higher monthly fees, with a higher drop-out rate concerning vocational education. However, they did not find evidence of an impact on the labor market (annual earnings, labor participation, the rate of stable jobs, and the intensive margin in the labor market). Bucarey et al. (2020) use a regression discontinuity design and the eligibility rules to examine the impact of Chile's student loan program on enrollment and find that student loans increase enrollment rates, particularly for students from low-income backgrounds. Finally, Card and Solis (2022) also use a discontinuity regression design and the sample of retakers, first-year students who re-applied to the test since the previous year they did not obtain the minimum score to qualify for the loan, to study the impact of student loans on no enrollment and degree completion. They find that, for higher-income students, the impacts on no enrollment decrease and are statistically insignificant from the third year onwards. Student loans increase degree completion rates by 2-4 percentage points. These studies provided the first evidence, with impact evaluation, that student loans can positively impact enrollment, retention, and course completion, particularly for disadvantaged students.

For Brazil, De Mello and Duarte (2020), exploring the expansion of Fies in 2010, the eligibility criteria of higher education institutions, and the difference-in-differences framework, found that Fies led to a 6% increase in the value of monthly fees, a lower price elasticity of demand, and an increase in the proportion of Fies students responsible for 56% of the net profit margin increase of educational institutions between 2010 and 2012. [4] Pontuschka (2016) also explored the eligibility criteria of higher education institutions and the differences-in-differences methodology with aggregation by the institution of data from the higher education student registration database (Higher Education Census). He identified an enrollment impact ranging from 6% to 9% (evasion results did not pass the robustness tests). Becker and Mendonça (2021), again with difference-in-difference and propensity score matching, in addition to the use of the database of the national high school student performance exam, which evaluates a sample of students at the beginning and end of the course, found that Fies increased the time students stayed connected to higher education institutions, around four months. Dobbin et al. (2021) applied a general equilibrium model to a state of the country (Rio de Janeiro), with an individualized and de-identified database but with de-identified data from the administrative record base for the management of Fies contracts. They found that Fies increased prices by 1.2% and enrollment by 11%. Rocha et al. (2016) used propensity score matching with difference-in-differences and crossed the administrative records of individualized and de-identified Fies

[4]This result is related to the program's design implemented between 2010 and 2014, with several incentive problems for higher education institutions and students. For example, the higher education institutions were not co-responsible for the retention of students in degrees or for default, and the students did not have co-participation in the current monthly fee (mechanisms that allow them to be informed of increases in monthly fees).

contracts with the formal labor market administrative records. They found that graduates with Fies earned an income 20% higher than non-higher education graduates.

There are limitations to findings from the student loan evaluation literature. The first problem is that some papers use registration information from a single institution or a subset of institutions, as already described by Solis (2017) and Card and Solis (2022). A database that centralizes all enrollment records in higher education is important. It allows the correction of attrition and errors in the classification of students applying for student loans, such as whether or not they enroll in higher education. The second problem is that the paper uses a restricted subsample of applicants for student loan programs, only those who have accessed higher education, because information on applicants for the student loan program who did not enter higher education is not available. The third problem is that most of the papers use the eligible and ineligible as an empirical strategy, given that they do not have identified administrative records of the student credit programs with the information of the students who applied for the student loan program, that is, who effectively moved due to the availability of the program (Rocha et al., 2016; Solis, 2017; Montoya et al., 2018; Bucarey et al., 2020; Becker and Mendonça, 2021; Dobbin et al., 2021). Card and Solis (2022) explain that the eventual comparison of eligible with ineligible around this score is problematic because the eligible can be induced by their eligibility and the consequent availability of credit, as recurrently explored (Solis, 2017; Montoya et al., 2018; Bucarey et al., 2020; Dobbin et al., 2021; De Mello and Duarte, 2020). The hypothesis that comparing eligible and non-eligible people around the rule cutoff solves the self-selection problem is strong when one does not consider how they moved and whether or not they applied for the student loan. Card and Solis (2022) solved this problem in their evaluation of Chile's program by using data from retakers, who they noted had applied for student loan access during the second year of the degree. Faced with this way of solving the self-selection problem, they point out that the work contributes marginally to the literature on the effects of student credit, as they analyzed the impacts on a very specific student population - the retakers are a very small portion of the eligible group and with much lower performance in the standardized test used for eligibility for the program and the high school test.

The literature with individualized data and with enrollments throughout the higher education system is recent (Bucarey et al., 2020; Montoya et al., 2018; Dobbin et al., 2021; Card and Solis, 2022), but she does not use administrative information on student loan applicants and, therefore, do not resolve the existing selection bias by using eventual students who do not need the student loan as a control. This paper contributes to the literature due to the richness of the data used, which explores data centered on a panel of students who applied in the Fies selection processes between the second semester of 2015 and the second semester of 2017, with information about them before this enrollment and

throughout of higher education across the country. I based the sample used in this article on all applicants to the program in the analyzed period and crossed it with data from all enrollments in higher education. As a result, the base reaches a wide range of the Brazilian population, with a more representative and heterogeneous profile since the criteria for accessing the program are broad. This database allows for monitoring non-higher education students interested in enrolling, providing a more direct way of analyzing access and the persistence of those enrolled without restricting the control sample only to those enrolled. This aspect is an important consideration for evaluating persistence in higher education.

My main contribution is that this work proposes to robustly assess the impacts of loan students on educational indicators from the crossing of databases at the individual and identified level and the methodology of regression in discontinuity, possible to be explored with the alterations that promoted competition for Fies vacancies in several courses, which took place in mid-2015. Crossing the databases of the National High School Examination (Enem), the Fies Selection, the Fies Computerized System (Sisfies), and the Higher Education Census make it possible to verify the balance in variables such as ability in that the student's performance in the Enem serves as a proxy for their ability before admission. The discontinuity regression approach has high internal validity since the hypothesis of independence between access to the program and the observable and non-observable variables is less strong than in previous papers, which used an eligibility rule to determine the control group. I achieve greater external validity in my work because I have a broad set of applicants to the student loan program. The results show that the Fies positively and significantly impact access to higher education and the student's preferred degrees. This result occurs in the year of application in the Fies selection process and the following two years. The impact of enrolling in the degree is positive and occurs only after 3 or 4 years of application in the selection process, that is, towards the end of the degree. There is no impact on course migration, a possible sign that students enter the degree they want. The results on the impact on the degree completion time of graduating in the expected regular time are negative. Still, they are not robust to changes in kernel and order of the polynomial.

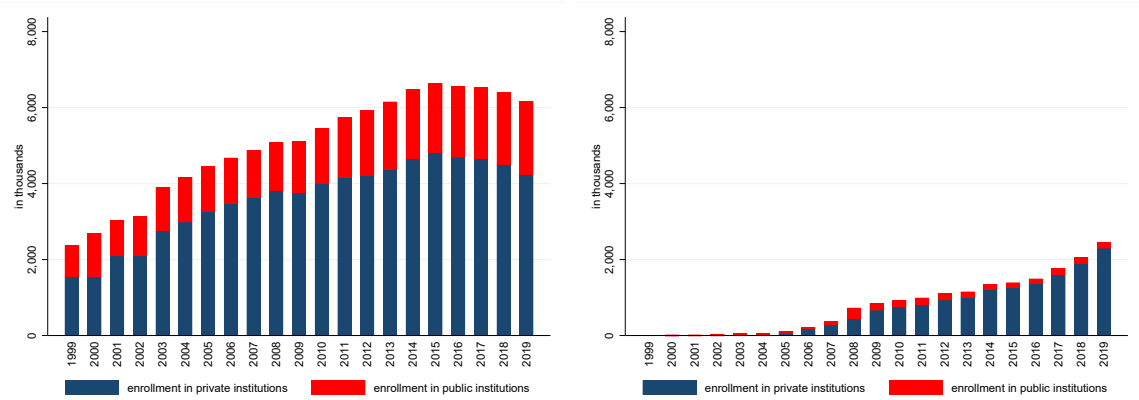
The paper is organized as follows. The next section presents the institutional background, with the antecedents of higher education in Brazil and the Fies. In the third section, I describe the databases and administrative records. The fourth contains the details of the empirical strategy. The fifth and sixth sections show the main results and robustness tests. Finally, I describe the conclusions and policy recommendations from this paper.

2 Institutional background

2.1 Higher Education in Brazil

The higher education system in Brazil is composed of public institutions, managed by the federal government or subnational entities (in general, by state governments), and private institutions, whether for profit or not. Public institutions have no monthly charge, as the Federal Constitution prohibits this, and access to degrees occurs through generally competitive selection processes. On the other hand, private institutions charge monthly fees that vary according to the degrees and the reputation of the institution. Access is known to be relatively easy, except for a few institutions in the main capitals and a few degrees, for example, graduation in medicine.

Figure 1: Evolution of enrollments in face-to-face and distance graduations in public and private institutions



(a) Face-to-face

(b) Distance

Figure 1 shows the total number of enrollments in higher education undergraduate degrees in private and public institutions, broken down by face-to-face (a) and distance learning (b). Enrollments in face-to-face courses at private institutions have accounted for more than 70% of education enrollment since 2003. The increase in enrollment in the private network took place even after efforts, mainly by the federal government, to expand public institutions.[5] In distance enrollments, in 2002, only 15% were in private institutions, rising to more than 80% since 2010. Distance courses significantly increased over time, but most enrollments are in face-to-face courses at private institutions. These data show that the debate on access policies to higher education needs to consider enrollments in private institutions with a special focus.

[5]For example, in 2007, the federal government implemented the Support Program for Restructuring and Expansion Plans for Federal Universities (Reuni), aiming to double the number of students in undergraduate courses in ten years, starting in 2008. Between 2008 and 2017, there was an increase of around 87% in federal higher education institutions enrollments.

2.2 Brazilian federal government student loan - Fies

The Student Financing Fund (in Portuguese, *Fundo de Financiamento Estudantil* - Fies) is a student loan program designed to finance tuition fees[6], instituted by the federal government in 1999 to expand access to private higher education. In general, students who have not had access to public educational institutions or have not had access to scholarship programs (non-refundable funding)[7] seek student loans. Reformulations in Fies occurred several times, and until the period of our analysis, there were three main designs. We explored the changes in the program in the last design described to evaluate the impact of Fies on higher education indicators.

The first design was in force until 2009, when Fies had restricted access, with more bureaucratic contractual conditions, such as the presentation of a guarantor and the requirement that the student and guarantor reputable and transparent application (not included in the registration records of defaulting debtors) when signing the contract and its amendment. It was also a little attractive, such as high-interest rates[8] and financing of a maximum of 50% of the tuition fee by educational institutions.[9] The second design was in force from 2010 to 2014. There was a reduction in the effective interest rate to 3.4% p.a., the financing of up to 100% of the tuition fee, and the release of students with a family income of up to 1.5 minimum wages to present a guarantor, with the institution of the Educational Credit Operations Guarantee Fund (in Portuguese *Fundo de Garantia de Operações de Crédito Educativo* - Fgeduc) to guarantee up to 90% of the value of these students' contracts. Finally, the third design was implemented between the second half of 2015 and the end of 2017; the new contracts now have effective interest rates of 6.5% p.a.[10], there was continuity in the percentage of up to 100% of the financing, and the amortization takes place up to 3 times the period of the financed degree.[11]

In the analyzed period, the following formula determined the percentage of tuition financing

[6]As in Chile, the loans do not cover living expenses or any other expenses associated with attending college (books, transportation, etc.).

[7]The main financial aid program of the federal government, created in 2004, to help with full and partial tuition fees (50% of the monthly fee covered) is the University for All Program (Prouni). Students with partial scholarships can apply for funding from Fies. In the analyzed period, x% of those hired with Fie had partial scholarships from Prouni.

[8]The effective interest rate on the financing ranged from 9% p.a. in 2006 to 6.5% p.a. between 2006 and 2009, except for undergraduate degrees, pedagogy, higher education, and higher technology courses, which had a rate of 3.5% p.a..

[9]Except for priority degrees (chemistry, physics, mathematics, biology, engineering, medicine, geology, and higher education in technology), where Fies covered up to 75% of these values.

[10]This interest rate had a high implicit subsidy. According to monthly debt reports published by the National Treasury Secretariat, in August 2015, the average cost of the federal public debt was 15.93% p.a.. This cost was 10.29% p.a. in December 2017.

[11]In 2018, the government implemented another design, but it is beyond the scope of this work.

obtained by those enrolled in the selection process:

$$f_{ij} = \left[1 - \left(\frac{[k_{ij}R_i] - d_i}{m}\right)\right] 100 \,, \tag{1}$$

Where f_ij is the percentage of financing of the fee charged of the student i with income family bracket j, k_ij is the percentage of marginal commitment of with income family bracket j of the student i, R_i is the per capita gross monthly family income in reais of the student i, d_i is the portion to be deducted by per capita gross monthly family income of the student i, m is the value of the educational fee charged by the HEI in reais. The first range is up to 0.5 minimum wages per capita brutal monthly family income; for this range, k equals 15%. The income ranges that determine k vary by 0.5 minimum wages per capita brutal monthly family income, and the value of k varies by 11.5% for each K range. Thus, the next range is from 0.5 to 1.0 minimum wage, and the value of k is equal to 15% + 11.5%; after 1.0 to 1.5 minimum wages where k is equal to $15\% + (11.5\% x 2)$, and so on up to 2.5 minimum wages. Along with applying this formula, there was the rule that the amount paid by the student per month of co-payment, $1 - f$, must be at least R$50.00.

Until 2009 there was no eligibility restriction based on family income. However, between 2010 and 2014, this criterion was gross monthly family income of up to 20 minimum wages, changing to per capita family income of up to 2.5 minimum wages in 2015 and up to 3 minimum wages in other years. Given the country's income distribution, the program has always been wide-ranging. For example, from 2013 to 2017, more than 90% of people over 17 with complete high school or incomplete higher education could apply for Fies.[12] It is important to highlight this point because the eligibility criterion based on family income does not restrict the external validity of our results. The other existing eligibility criterion refers to the quality of degrees and institutions implemented in 2010 when the government established that degrees or educational institutions must meet a minimum quality parameter, score greater than or equal to 3 in the National Higher Education Assessment System (Sinaes) to be eligible for Fies. Sinaes consists of three assessments: Enade evaluates the performance of selected students in the first and last year of the course; Preliminary Course Concept (CPC) considers the value added by the course to concluding students, based on their performance in Enade and Enem, faculty, and student's perception of their formative process; and General Index of Courses IGC is an average of undergraduate courses' performance at CPC, weighted by enrollments, and performance of master's and doctoral courses in CAPES grade, also weighted by enrollments.[13] This criterion is also not restrictive. For example, in 2015 and 2017, around 85% of private

[12]Data were obtained from the IBGE's Continuous National Sample Survey by Households (PNAD-Continuous).

[13]The reference three-year period is the basis for calculating the IGC. These three assessments range from 0 to 5.

institutions evaluated in the IGC were eligible for Fies.

In 2015, the Fies selection process started to use the Enem performance in two ways. The Enem, described in detail later, is a standardized test measuring students' academic ability and under the management of National Institute of Educational Studies and Research Anísio Teixeira (in Portuguese, *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira* - selection), an agency linked to the Ministry of Education. The first was the minimum score criterion to be eligible, instituted in December 2014. High school graduates from 2010[14] onwards had to obtain an arithmetic average of grades in the Enem tests (reading and codes, mathematics, natural sciences, and human sciences) equal to or greater than 450 points and a score of essay other than zero to be eligible for Fies.[15] It is important to note that since 2010, the program required participation in the Enem as a prerequisite for the application process at Fies that students who had completed high school as of 2010, both for incoming students and for students already enrolled in higher education. However, during this period, the program did not require any minimum performance criteria in the exam. As the requirement to participate in the Enem was prior to 2015, the change that allowed exploring the performance in the Enem did not translate into a new cost for Fies applicants.[16] The eligibility score of 450 points in Enem is not very restrictive, as shown in Figure 2. This eligibility score is at the end of the distribution curve for students' scores enrolled in public institutions (3(a)). Also is below the average of students in private institutions (3(b)), including those that adhered to Fies (3(c)). This comparison is valid before and after the implantation 2015 rule, which is important for discussing the external validity of the results of this work.

The second use of Enem is that, due to the drastic drop in the number of vacancies offered motivated by fiscal restrictions, from the second half of 2015, student access through a selection process based on ranking in the Enem. The dispute for places established a passing grade in the most demanded degree, generating an exogenous variation around these cutoff points. The student requests access to the program only through a competitive selection process at the beginning of the first and second elective semesters.[17] Thus, the competition for available places in Fies in degrees of greater demand begins, in which the

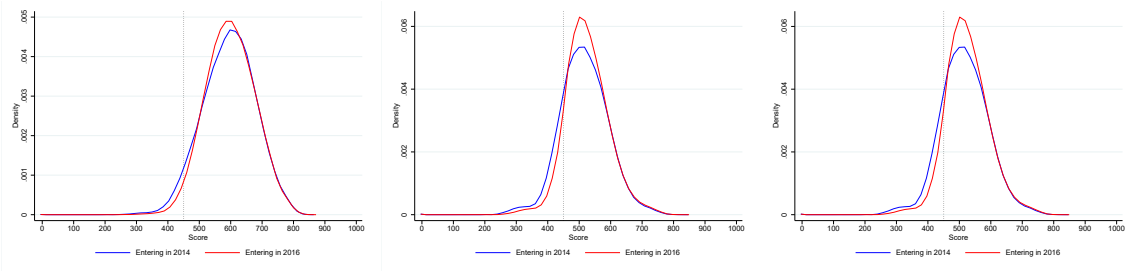
[14]Those who completed high school before 2010 were exempt from this obligation if they could prove this condition to the Permanent Commission for Supervision and Monitoring (CPSA), linked to the higher education institution.

[15]Normative Ordinance of the Ministry of Education Nº 21, December 2014. More details in: https://sisfiesportal.mec.gov.br/arquivos/portaria_normativa_21_26122014_compilada_050115.pdf.

[16]The enrollment fee for Enem was R$35.00 in 2010 (6.0% of the minimum wage in force for the year), and in 2017, it was R$82.14 (8.8% of the minimum wage in force for the year). Low-income students can apply for exemption after filling out forms and submitting documentation to selection. From 2009 to 2016, the test takes place in 2 afternoon periods on the weekend.

[17]Before, the student could apply for Fies at any time and, there was no dispute over the vacancies offered.

Figure 2: Distribution of scores in the Enem

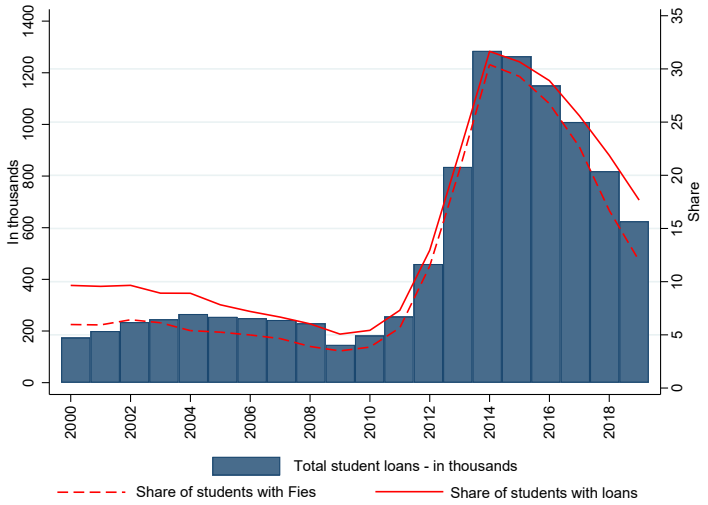


(a) Public institutions (b) Private institutions (c) Private institutions with Fies

Source: Enem Administrative Records, Census of Higher Education, and Fies administrative records. Notes: Vertical dashed line represents the 450-point eligibility cutoff in the Enem. The score in the Enem is calculated by the arithmetic mean of the grades in: i) mathematics; ii) languages and codes; iii) human sciences; iv) natural sciences; and v) essay. Graph (a) considers the score in the Enem taken in the t-1 year because when the public universities use this exam for admission in t, it is this score that is valid. For the purposes of comparing students' academic ability, graphs (b) and (c) follow the same logic. Graph (c) show score in private colleges that have students with Fies. The graphs compared the scores of students who enrolled in the degree in the census year and not in the entry year because the Fies student can enter the program in any semester of the course.

program selects students with the highest average grade in the objective test of Enem.[18] In this new design, meeting the eligibility criteria no longer guarantees access to the program. When applying in Fies, the student knows his Enem average score and, consequently, if his score in Enem meets the eligibility requirement of 450 points and if the essay grade was greater than zero. However, they do not know if they will be selected because they are unaware of the demand for the degree at the applied institution and the score in the Enem of their competitors. This exogeneity provides a good candidate for a control group: students who applied to the program and scored on Enem just below the passing grade.

Figure 3: Student loan enrollment

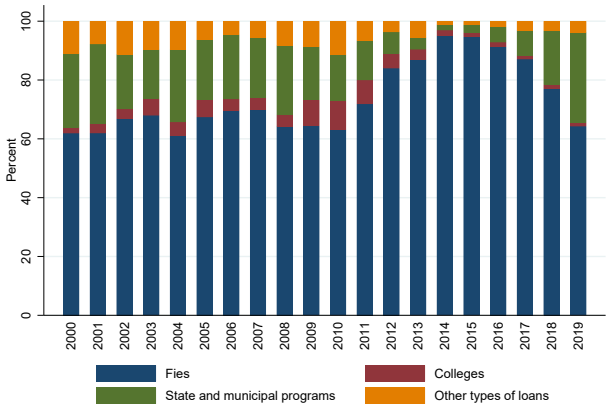


Source: Census of Higher Education. Note: Total student loans include loans offered by federal, state, and municipal governments, private higher education institutions, and external entities such as private banks. The share of enrollment with loans represents the proportion of students in private institutions with any student loan.

[18]Normative Ordinance of the Ministry of Education Nº 8, July 2015. More details in: http://portalfies.mec.gov.br/arquivos/portaria_8_2_07_2015_2-2015.pdf.

As shown in Figure 3, there was relative stability in the number of students in the program until 2010, from 2011 to 2014, a strong expansion, and from 2015 onwards, a sharp drop in the number of students. At the beginning of the program, only 200,000 students had a student loan, representing about 10% of all private enrollments in higher education, of which half were already from Fies. Therefore, since its inception, Fies has been the main form of access to student loans. Figure 4 shows the importance of Fies compared to other types of student loans. When Fies is a more modest program, we see greater participation in other student loans, such as those offered by other federal entities (states and municipalities), higher education institutions, or other private institutions. Nevertheless, the share of Fies in the total of students financed by the student loan is always above 60%, from 2012 to 2017 above 80%, with the peak in 2014 and 2015 when around 95% of students with student loans were from Fies. In the period we analyzed, more than 85% of students with student loans were from Fies.

Figure 4: Distribution of student loans by type



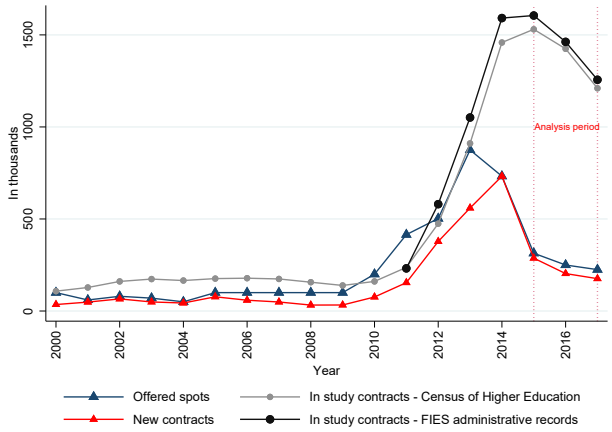
Source: Census of Higher Education. Note: Other types of loans include student loans offered by startups and private banks.

Figure 5 shows the number of vacancies offered, new contracts, and, according to data from the Higher Education Census and the Fies administrative record - Fies Informatized System (in Portuguese, *Sistema Informatizado do Fies* - Sisfies), the number of contracts in the financing phase (when students are still in higher education).[19] The government reduced this availability from 2014 onwards due to the fiscal restriction framework. In

[19]According to the data source, there is a difference between the contracts in the financing phase because the different bodies managing these databases do not share and cross-reference the information in their administrative records. There are two hypotheses for this: 1) there is a failure in the administrative record of the Higher Education Census, which did not include these students, and 2) there may be leakage of the policy, in which students receive the benefit but do not use it. Between 2013 and 2017, we identified 204,218 students on the Fies administrative record but not on the Higher Education Census record. However, as will be discussed in more detail in the data and empirical strategy sections, this will not be a problem for the analysis.

the analyzed design, there was a drastic drop in vacancies and the total number of new contracts signed, averaging 263,000 and 222,000, respectively. It is evident that from 2014 onwards, there was an approximation between the offer of new vacancies in the program and their occupation, given by the new contracts signed, since before 2014, there were leftover vacancies available in the Fies. This greater occupation of available vacancies is related to the change also explored in the empirical strategy, which is the existence of a selection process for the occupation of vacancies available in the program.

Figure 5: Offered spots and contracts signed at Fies - 2000-2017



Source: Ministry of Education, Census of Higher Education, and Enem administrative records. Note: As no administrative data is crossing between the agencies of the Ministry of Education that manage the Higher Education Census and the Fies, there is a discrepancy between the information about the number of students in study contracts in their respective databases.

The access process to Fies consists of a few steps. First, the higher education institutions present to the Ministry of Education a proposal to offer vacancies. The vacancies offered in the program are selected by the Ministry of Education based mainly on available resources, social relevance determined by micro-region[20], priority degrees[21], and prioritization of graduations and institutions with the highest score in Sinaes.[22] Then, students apply for free in the program's selection processes at the beginning of the first and second semesters,

[20]To define relevance, the Education Ministry considers the following information: i) demand for higher education, calculated based on data from the National Secondary Education Examination (Enem); ii) demand for student loans, calculated based on Fies data from the previous year; and iii) Municipal Human Development Index (HDIM) of the micro-region, calculated from the average HDI of the municipalities that comprise it, according to studies developed by the Program of the United Nations for Development Brazil (PNUD/Brazil) and by research institutes in the country.

[21]Up to 60% of vacancies per micro-region are allocated to priority degrees, with up to 30% for courses in the health area, up to 24% for engineering and computer science degrees, and the remaining 6% for undergraduate degrees, pedagogy and higher education degrees for train teachers able to teach in kindergarten and the first years of elementary school.

[22]There was a prioritization in the selection of spots for degrees based on their concept score: up to fifty percent of the spots were reserved for concept five degrees, up to forty percent for concept four degrees, up to thirty percent for concept three degrees, and up to twenty-five percent for degrees that have only been authorized by regulatory acts.

usually in February and June, through the Fies Selection (in Portuguese, *Fies Seleção*). The system verifies that applicants meet the minimum eligibility score requirements and ranks the students who are competing for the same vacancies by the score on the Enem. Thus, students who obtained a higher score in the Enem than their competition are pre-selected. Each degree of the institutions in Fies has a first passing grade. Some students are pre-selected in the calls and give up hiring Fies. With this, the students on the waiting list are successively called and receive the pre-selected classification. Therefore, there is a final passing grade. Students selected by these cutoffs go to the bank branch, defined by the federal government as the financial agent of the program, to formalize their financing agreement and, every six months, for its amendment.[23]

On this, we observed that in the analyzed period, on average, in the courses of students with Fies last nine semesters, about 68.4% of the students joined at the beginning of the degree (they still haven't attended any semester), and of those who joined Fies during the course, on average, took about two and a half semesters. The ten degrees with the highest number of students financed are Law, Nursing, Civil Engineering, Psychology, Administration, Physiotherapy, Nutrition, Dentistry, Accounting, and Pharmacy. These degrees accounted for 59.4% of student loan agreements in the period. On average, students requested funding of 81% (sd 14%) of the degree semester value and obtained 74% (sd 23.7%) of funding.[24]

2.3 National High School Exam

The National High School Examination (in Portuguese, *Exame Nacional do Ensino Médio* - Enem) is a standardized test created in 1998 to evaluate the school performance of students who graduated from high school.[25] The Inep is responsible for its application throughout the country. Initially, it took place in one day on the weekend, contained 63 questions and one essay. Still, since 2009 it has taken place over two days on weekends in late October or early November, covering 180 questions divided into blocks of 45 questions for each area of the knowledge (mathematics, languages and codes, human sciences, and natural sciences) - and one more essay. In addition, as of 2009, Enem started to follow

[23]The loan payment to institutions is made monthly through the issuance of non-transferable public securities by the National Treasury, used to pay debts and taxes with the federal government, or, if these possibilities are exhausted, higher education institutions may periodically redeem these titles.

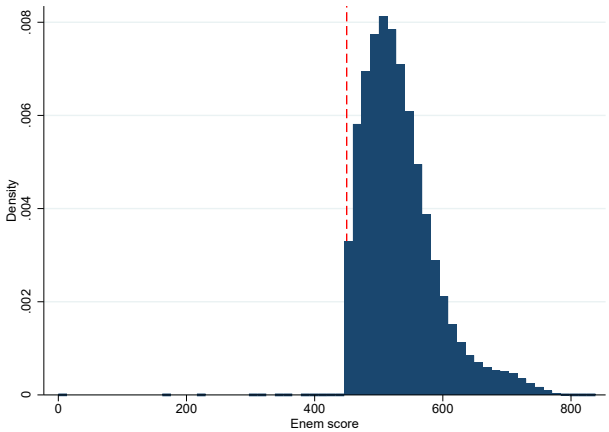
[24]The average nominal values of the semesters financed by students entering the program were R$ 6,572.1 (sd 4,652.9) in the second half of 2015, R$ 6,502.9 (sd 4,456.7) in the first half of 2016, R$ 6,749.2 (sd 4,757.1) in the second half of 2016, from R$ 7,252.4 (sd 5,087.5) in the first half of 2017 and R$ 7,170.3 (sd 4,885.3) in the second half of 2017.

[25]The government instituted the exam through the Ordinance of the Ministry of Education Nº 438 of May 1998. According to the Ordinance, the objectives of this exam are: "i) to provide citizens with parameters for self-assessment, to continue their training and their insertion in the labor market; ii) create a national reference for graduates from any of the secondary education modalities; iii) provide subsidies to the different modalities of access to higher education; and, iv) constitute a modality of access to post-secondary vocational degrees."

the Item Response Theory (IRT), allowing for a comparison of the applicant's academic ability over the years.

The Enem is a test with a wide territorial reach. During the period, there was growth in the application in municipalities located in the north, northeast, and midwest, regions with poorer municipalities and greater mobility problems (Figure A.1 in Appendix **??** shows the municipalities that applied the Enem in 2010 and 2016). This signals the breadth of access to the exam used in the Fies selection process. Figure 6 shows the grades of those applicants in the Fies selection processes in the analyzed period.

Figure 6: Distribution of applicants' Enem scores



Source: Administrative database of the Fies selection processes. Note: Scores from applicants who participated in the selection processes during the analyzed period.

3 Database

3.1 Administrative database of the program's selection processes

The Selection Fies database derives from the program's selection system, administered by the Ministry of Education. The Fies selection process is similar to the one used in university entrance exams, where each educational institution that joined Fies offers vacancies by shift and degree. I accessed the data identified with the Individual Taxpayer Registration number[26], socioeconomic data, such as family and per capita income, and the institutions and degrees required by applicants in each Fies selection process. I know who the pre-selected applicants are in the first notice called through the information on the passing grade variable available in the base, the final passing grade, and all applicants pre-selected at the end of the process, and, consequently, the applicants who were not

[26]Inep masked the Individual Taxpayer Registration number with the unique code that allows correspondence with the other databases they managed, thus allowing the crossing between the databases used.

selected. I have data from the selection processes for the second half of 2015 and the second half of 2017. As those interested in Fies can apply again for the Fies selection process when they are not selected, I observe how many times they apply. I fix my sample in the first selection process in which he applies. On average, around 558,000 students applied in the selection processes in the first semester, and around 248,000 students applied in the selection processes in the second semester.

3.2 Student contract management database in the program funding phase

The National Education Development Fund (in Portuguese, Fundo Nacional de Desenvolvimento da Educação - FNDE), an autarchy linked to the Ministry of Education, manages the Fies Computerized System (in Portuguese, Sistema Informatizado do Fies - Sisfies). Sisfies manages contracts in the financing phase, i.e., when the student attends higher education with student financing. It is biannual because students amend contracts every six months; therefore, these amendments have contractual updates, such as the tuition fee for the degree. I accessed the identified database from this system between the second half of 2010 and the first half of 2021. The periods after the data I accessed from Fies Selection allow us to verify whether the applicants who participated in the selection processes in the analyzed period benefited from the program later. Students who accessed Fies after the analyzed period are excluded from the sample, as they would be classified as not receiving Fies when they actually would have received it.

With this database, I verified whether the students pre-selected in the Fies Selection had joined the program. The dummy about having or not contracted Fies comes from this database because there is an information error of 6.5% of students who appear in the Higher Census as having Fies but who do not appear in the Fies contracts base and of 4.2% of students who, according to the Higher Census, did not have Fies and who did. The available variables are the registration information of these students, such as occupation and address, and the contract information, such as degree, institution, and score in the National Higher Education Assessment System (Sinaes)[27], semester values

[27]Sinaes comprises three quality indicators for higher education degrees and institutions monitored by the Ministry of Education. The first is the National Performance Examination (Enade), which evaluates the performance of the contents of the degree of selected students in the first and last year of the degree. The second is the Preliminary Course Concept (CPC), which uses as parameters: i) an indicator of the value added by the degree to graduating students, based on their performance in the Enade and Enem; ii) some indicators related to the teachers (percentage of masters, doctors, and forms of working arrangements); and, iii) the student's perception of their training process in the degree. Finally, the third is the General Course Index (IGC), an average of the performance of undergraduate courses in the CPC weighted by enrollments in different courses, as well as the performance of master's and doctoral courses in the grade established by the Improvement Coordination Higher Education Personnel (CAPES), also weighted by enrollments in these courses.

with and without discount[28], the percentages of funding requested and approved. I also have information about the number of semesters of the degree the student attended when he joined the program, among other information.

3.3 Higher Education Census

The Higher Education Census compiles public and private higher education data in Brazil, with variables on students, educational institutions, degrees, and professors. Educational institutions annually fill out a questionnaire with the information requested for the Higher Education Census through an electronic system managed by Inep and through the accreditation and re-accreditation system of higher education degrees. As public and private educational institutions are subject to regulation by the Ministry of Education, filling out this questionnaire is mandatory. Another motivation for private institutions is that completing this system is a prerequisite for participating in federal government programs like Fies.

With this database, I set up a panel to monitor access, permanence, and graduation in higher education degrees by students who applied in the Fies selection process. I used data from the Higher Education Census between the years 2010 to 2021. The years before 2015, therefore before the design of the Fies studied, are important for the analysis of students who dropped out of the degree and may try to resume higher education with the program or take advantage of it to migrate degree, opting for a degree that was previously outside the budget constraint, for example.

My panel maintains observations by student and degree, regardless of class and higher education institution. On this basis, the degree code refers to a given degree class at a given institution. For example, if there are two classes in a law degree at a higher education institution on different shifts, each class receives a different degree code. I used a degree classification variable based on the International Standard Classification of Education – Fields of Education and Training (ISCED-F) of the United Nations Educational, Scientific, and Cultural Organization (Unesco) to make these degrees compatible. Thus, according to this degree classification variable, my observations accompany the higher-education student by degree. In addition, I have information on the modality the student attends (whether he attends the degree in person or at a distance), the shift (morning, afternoon, full-time, or evening), regular (or expected) time to complete the degree, the semester and year of entry, the year and the finishing semester, enrollment status, etc. With the variables of the degree's regular time and semesters of entry and conclusion, it is possible to define the variable of time of conclusion of the student's degree. I also have some

[28]According to the law, institutions must grant all benefits to other students without funding. However, there are no monitoring and control mechanisms and no way to verify whether institutions fulfill this obligation.

socioeconomic variables, such as age, race, and whether you attended high school in a public school.

3.4 Enem database

I set up a panel with data on Enem applicants from 2010 to 2016 since those selected in the last selection process of the analyzed period, the second half of 2017, may have, at the limit, used their score in the 2016 Enem. I have the scores on specific tests (reading and codes, mathematics, natural sciences, humanities, and essay) in the years in which the student took the Enem and additional socioeconomic information, such as the parent's education, the number of people living in the household and the family income (in intervals relative to the minimum wage, for example, between 2 and 5 minimum wages). I also checked how often students took the Enem until they entered Fies.

3.5 Matching and Final Dataset

In the analyzed period, 1,900,494 students enrolled in the Fies Selection. I initially reduced my base to 1,860,883 observations[29] due to consistency problems when crossing with the Superior Census. Additionally, there were some observations of students in the Sisfies database who do not appear in the Higher Education Census, totaling 23,995 students, and the base also was further reduced due to some flexibility in accepting students who had not taken the Enem until then when implementing the selection rule based on the Enem score. In the end, there was a loss of almost 2% from observations. The proportion of selected students is high (70.17%) because many are selected and do not hire, allowing the call list to be called. In my sample, 392,872 students were pre-selected and hired for funding in the selection process in which they participated.[30]

Since the information on enrollment status for the year is reported annually, the access variable to higher education or a degree is defined as access in the year of the Fies selection process and also as one or two years after application in the selection process. The main access results consider students not in higher education when they applied for Fies. For the enrollment variables, I consider the situation of active enrollment when in the year in some higher education degrees, the student was in the situation of active enrollment or as a graduate of the year. The migration variable took the value of 1 when the student's enrollment situation in higher education was locked, unlinked, or changed degree at the same institution in the previously enrolled degree, and up to 3 years after this situation, the student appears enrolled in a new degree.

[29] Some degree codes in the Fies Selection database do not have correspondence in the Higher Education Census, i.e., these degrees do not exist or need to be corrected in the Fies Selection database.

[30] It is worth noting that 215,343 students applied in the selection processes analyzed and were not selected but joined Fies through selection processes outside the analyzed period.

The graduate variable is defined using the degree completion variable, the degree's regular time, available in the Higher Education Census. This variable considers the year of admission and the time since degree completion. I also look at completing within a year or two of the expected year of completion. If the student appears with the registration situation as having graduated in the year expected for completion - graduate one year later or graduate two years later, this dummy receives the value of 1; otherwise, zero. Most degrees take about four or five years in the degree's regular time, but there are degrees with an expected time of 5 years.[31] As we only have data from the Higher Education Census up to 2021 and there are new students up to the second half of 2017, it's necessary to limit the sample to students who would be graduating within the period of the sample. Thus, for the result variable graduated in the expected year, these exercises consider students with degrees whose expected completion time is less than or equal to 2021. For the variable that graduated up to one or two years after the expected year, we limit the sample to students whose expected year of completion would be less than or equal to, respectively, 2020 and 2019.

Table A.1 from Appendix A.1 shows the descriptive statistics of Enem and Fies applicants. There are more black, brown, and indigenous people among Fies applicants, and their parents have less education. These students make the most effort to enter higher education because, on average, they take the Enem more often. Fies applicants had an average performance in Enem tests similar to Enem applicants.

4 Empirical strategy

With the competitive Fies selection process linked to the Enem score from the second half of 2015, there is an almost natural experiment. The passing grade in the Enem, which gives access to the different degrees of the educational institutions registered in the Fies, generates an exogenous variation that is as if there were randomization of funding close to this cutoff score. Selected students receive an offer of a percentage of financing for the degree, and they decide whether or not to take out the student loan. Some selected students give up taking out the loan, and the program selects new applicants to fill the available vacancies. In addition, some degrees receive numbers less than or equal to the number of places available. Thus, I adopt the fuzzy Regression Discontinuity Design (RDD) as an empirical strategy (Lee and Lemieux, 2010) to estimate Fies' impact on access to and permanence in higher education.

[31]The percentage of students in degrees with a regular duration of fewer than four years, whether we consider students when they applied to Fies for the first time, is 5% and 4.7% without this filter of being in the degree that applied for the first time turn. Of the students who graduated in our sample, the average completion time is 4.35 years.

4.1 Identification hypothesis

The exogeneity of the degree passing grades stems from the fact that Fies students cannot manipulate their Enem score or the cutoff score for the degrees registered in Fies. Enem has strict control policies in which the makers of the objective and essay tests are kept anonymous and confidential. The objective test consists of multiple-choice questions drawn randomly from a bank of questions developed over the years. A photo-optical device grades this objective test. For the essay correction, the hired proofreaders and the information of the students who will have their essays corrected are kept anonymous and confidential.

The passing grade for degree in Fies depends on the number of vacancies made available by educational institutions and the demand for these degree in Fies, so the cutoff score is defined only at the end of the selection process for the program centralized at the Ministry of Education. As previously mentioned, the Enem is applied annually between the end of October and the beginning of November of year t. The selection process begins at the beginning of the first and second semesters of year t+1. An important point for the validity of the explored exogeneity is that the threshold is unknown during the selection process (Zimmerman, 2014).

Students applied in the Fies selection processes are selected in the program if and only if:

$$FIES_i = \begin{cases} 1 & \text{if} \quad x_i \geq pg_d \\ 0 & \text{if} \quad x_i < pg_d \end{cases} \tag{2}$$

where $FIES_i$ is a *dummy* that represents the student's access to Fies of the student i, $x_i \in [0, 1000]$ is his Enem score and pg_d is the cutoff north for the selection in the degrees d of the program. Therefore, in (2) we see that students who apply for Fies receive funding if their Enem score is greater than or equal to the passing grade in the applied degree, that is, the treatment is a function of this score. Similar to what happens in college entrance exams, we have a first passing grade and a final passing grade, because some selected students give up taking out the student loan, as already mentioned before.

As my sample is of those applied in the program's selection process, therefore, students that I observe who moved to apply for the student loan, I control by definition the variable that in previous works is unobservable: effective interest in the program. The identification hypothesis is that, for those interested in the program, the potential results of taking or not taking out student loans are the same for individuals who are on the right and left of the cutoff - the degree passing grade. This identification hypothesis can be represented by:

$$(Y(0), Y(1)) \perp FIES | X = pg_d \tag{3}$$

where $Y(0)$ and $Y(1)$ are, respectively, the result for the same individual of contracting and not contracting Fies. The score on the Enem is the *forcing variable* because it reproduces the results of randomization in $FIES$ when the score on the Enem is very close to the final passing grade of the degree. Thus, the group of students applying for Fies to the left of the passing grade becomes the counterfactual for students applying for Fies to the right of the final passing grade. With my program data, both treated and control groups applied for Fies.

There are two cutoff marks in my database. The first is the first passing grade, defined as the initial ranking of the applicants' Enem scores up to the number of places offered in the program. However, some selected applicants do not contract Fies, either because they did not submit the required documentation or because they did not find the contract conditions attractive (for example, they were not interested in the percentage of financing offered by the program), there are new calls from applicants who were on the lists of waiting, as occurs in college selection processes. Thus, a second cutoff score is available in the database: the final passing grade, defined after running these waiting lists. The final passing grade is the cutoff point presented in the main results since the monotonicity hypothesis is weaker since applicants with a score above this cutoff point hire Fies.

4.2 Estimation model

Estimates from local linear regression methods are attractive due to their properties of generating robust border estimates (Imbens and Kalyanaraman, 2012). The first model to be estimated assumes additive and linear specification in reduced form:

$$Y_i = \beta_0 + \beta_1 x_i \mathbb{1}(x_i \geq pg_d) + \epsilon_i \tag{4}$$

where Y_i represents the explored result variables, β_1 is the parameter of interest - local average treatment effects (LATE) of Fies, and x_i and pg_d are, respectively, the relative score in the Enem and the degree final passing grade. The estimates models are generated with the procedures established by (Calonico et al., 2014, 2017), using the optimal bandwidth selection and robust standard errors. This package generates consistent estimates of the RDD, as it properly and robustly corrects problems such as unobserved heterogeneity, measurement errors, autocorrelation, and endogeneity. The models also use the uniform kernel and the polynomial of degree 1 in the main models, following (Gelman and Imbens, 2019) who show that high-order polynomials can generate extremely flexible curves that fit the data very well, but do not necessarily reflect the true relationship between the treatment variable and the outcome. The use of high-order polynomials can result in overly smooth and inaccurate estimates of causal effects, leading to misleading conclusions.

In addition, models with selected ad hoc bandwidth of 30 points, with changes in the kernel to triangulate together with a polynomial of degree 2, besides estimates with covariates, are presented. All these models run again with clustering in the standard error at the shift (morning, afternoon, or evening), degree, higher education institution, and selection process corresponding to the level of selection from the applicants, following recent clustering literature, which demonstrates that the appropriate cluster level follows the level of treatment assignment (Abadie et al., 2020, 2023). The main covariates are dummies of race, gender, singleness, father's and mother's schooling, the fixed effect of the selection process, and degree competition, among others.

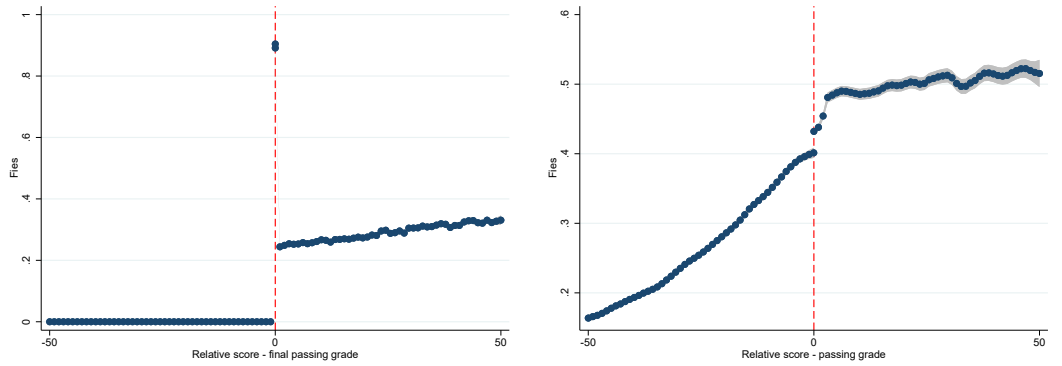
To evaluate the impact of Fies on access to higher education, I defined the first result variable as the dummy for access to higher education, which is equal to 1 if the student who applied for Fies accessed higher education in the year in which he participated in the process selective, if accessed within one or two years after this selection process. I also evaluated whether access to higher education occurs in the Fies applicant's degree of interest, setting this dummy variable equal to 1 if he enrolled in the degree he requested in the selection process since applicants who did not benefit from the program may, for example, enrolling in a degree that has the cheapest monthly fee, due to current income constraints. To evaluate the impact on permanence, I defined the dummy variable of enrolling in higher education one, two, three, or four years after the selection process. As with Montoya, Solis, and Card, I use the term enrollment (or non-enrollment, which would be equal to "dropout status") to avoid confusion with "event dropout". The situation observed in the year is not conditional on the situation in the previous year. As students can change degrees over the period they are in higher education, I also analyzed the effect of Fies on degree migration, where the dummy degree migration assumes the value of 1 if the student enters a new degree in the following year or within two years after leaving or withdrawing from the previous degree, that is, without having completed that previous degree and so that the year of admission to the new degree is greater than or equal to the year of the previous degree. Finally, I evaluate the impact of Fies on degree completion if it is completed within the expected time, using the variable degree's regular time, explained before. I also observed whether Fies students graduate up to one or two years beyond the expected time.

4.3 Validity of the RDD

Figure 7 shows the discontinuity in access to Fies from the final passing grade. There is an accumulation of observations exactly in the passing grade because, by definition, all students with the same grade as the passing grade hire the Fies.

Following the procedures proposed by (McCrary, 2008), in Figure 8 I estimated the

Figure 7: Score in Enem and access to Fies for applicants in the selection process

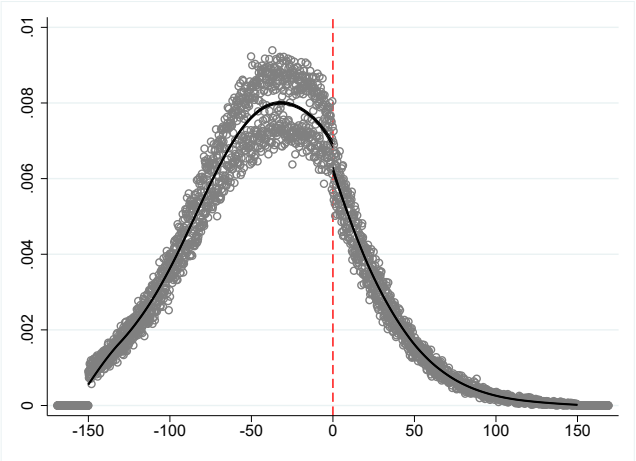


(a) Final passing grade (b) First passing grade

Source: Administrative database of the program's selection processes and Student contract management database in the program funding phase. Notes: The vertical dashed line represents the Fies cutoff score, where graph (a) uses the final cutoff score, defined after waiting lists, and graph (b) uses the first cutoff score, defined by the initial ranking of the applicants' Enem score up to the number of vacancies offered. One bin corresponds to one point on the Enem. Relative grade is equal to the grade on the Enem minus the cut-off grade used. Result with 95% confidence interval.

polynomials for the frequency distribution of Enem scores close to the final passing grade, used in our main estimations. The attached table A.2 shows the RD estimates for the covariates. In general, there are no discontinuities in the covariates. The exception occurs in variables such as Father with a postgraduate degree, the number of times applied for Fies, per capita family income, and competition for vacancies (number of applicants per vacancy). We estimate the RD models with controls and show no large variations in the estimates obtained.

Figure 8: Test for smooth histogram of Enem scores (McCrary Test)



Source: Administrative database of the program's selection processes and Student contract management database in the program funding phase. Note: The vertical line at 0 corresponds to the final cutoff grade, and the relative grades correspond to the average Enem score minus the applied degree final passing grade.

5 Results

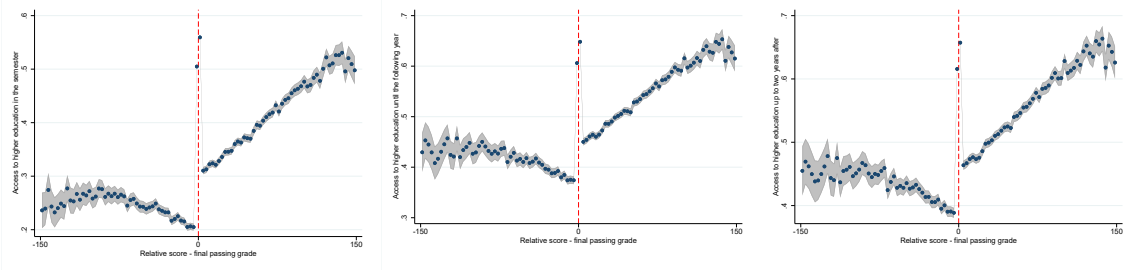
Below we show our main results, considering the final passing grade. Each table presents eight results, the first four without clustering (columns I to IV) and the last 4 (columns V to VIII) clustered at the level of the higher education institution, course, semester of the selection process and shift, which is the level where students are selected in the Fies selection process. Within this subset of 4 results, the first columns (columns I and V) show the results with a polynomial of degree 1, uniform kernel, and optimal bandwidth. The second columns (columns II and VI) consider the same polynomial and kernel but with an ad hoc bandwidth of 30 points. The third columns (columns III and VII) show the results with the same polynomial and kernel but with optimal bandwidth and controls. Finally, the last columns (IV and VIII) modify the polynomial to degree 2 and the kernel to triangular, including the optimal bandwidths and the controls. The controls used are the father's and mother's schooling, race dummies, single, whether he studied high school in a public school, whether he attended regular high school and the selection process, age, times he took the Enem, number of residents in the household, vacancy competition indicator (ratio between applicants and vacancies), the value of per capita family income and semester value. All estimated results using the (Calonico et al., 2014, 2017) package and the graphs that illustrate the effect of Fies on the variables of interest follow the recent findings of (Korting et al., 2021).

5.1 Access to higher education and degrees of interest

Figure 9 illustrates the effects of Fies on access to higher education for students who were out of higher education when they applied for the program. As shown by Table 1, the effects of Fies on access to higher education are high, statistically significant, and robust to the classic changes used - variations in bandwidth, polynomial, kernel, and inclusion of controls. The effect is greater when I consider the semester in which applicants provided the Fies; it drops when considering a period of up to 2 years after the semester of application to Fies. The point estimate suggests that students above the final passing grade have a chance of entering higher education in the semester of applying for the Fies of 47.7 percentage points more than students below that threshold. After two years, some students below this cutoff score enter higher education even without the Fies. Still, even so, students above the threshold have a chance of 31.5 percentage points more of entering higher education.

With the previous result, a question can be established: does the Fies access result remain if I consider the impact on access to the degrees applied in the selection process? This question is interesting because the Fies can be an opportunity for students to access degrees of greater interest; for example, due to budget constraints, they can be in higher

Figure 9: RD Plot - Access to higher education for those outside higher education with final passing grade



(a) In the semester (b) Up to one year after (c) Up to two years after

Source: Administrative database of the Fies selection processes, Student contract management database in the program funding phase and Higher Education Census.
Note: Vertical line represents the cutoff score, where the Enem scores on the horizontal axis is about the cutoff score given by the last call of the program's selection process. Estimates obtained with package from (Calonico et al., 2017), with the polynomial of degree 1 and uniform kernel. Each dot represents the number of students in a 3-point bin of Enem scores. The shaded region shows 95% confidence intervals.

Table 1: Access to higher education for those outside higher education - final passing grade

						Cluster-robust		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Access in the semester				
	0.584***	0.572***	0.514***	0.476***	0.582***	0.57***	0.512***	0.477***
	(0.029)	(0.027)	(0.048)	(0.067)	(0.031)	(0.028)	(0.049)	(0.068)
Bandwith	16.683	30	20.271	48.823	16.601	30	19.786	48.968
n	206,320	333,206	74,287	139,811	204,348	331,164	73,194	140,119
				Access up to one year after				
	0.432***	0.442***	0.316***	0.312***	0.425***	0.438***	0.315***	0.312***
	(0.035)	(0.03)	(0.047)	(0.058)	(0.036)	(0.031)	(0.048)	(0.059)
Bandwidth	14.935	30	19.838	54.764	14.579	30	19.584	54.275
n	188,532	333,206	73,284	152,295	183,817	331,164	72,738	151,218
				Access up to two years after				
	0.425***	0.438***	0.313***	0.315***	0.423***	0.433***	0.315***	0.315***
	(0.036)	(0.03)	(0.047)	(0.07)	(0.037)	(0.031)	(0.048)	(0.07)
Bandwidth	14.298	30	19.668	43.922	14.281	30	19.342	43.92
n	181,969	333,206	72,907	129,240	180,924	331,164	72,174	129,177
N	842,205	842,205	255,540	255,540	838,201	838,201	255,540	255,540
OP	1th	1th	1th	2th	1th	1th	1th	2th
Kernel	Uniform	Uniform	Uniform	Triangular	Uniform	Uniform	Uniform	Triangular
Covariates	No	No	Yes	Yes	No	No	Yes	Yes

Note: The sample used only considers students who, between 2010 and the year of enrollment in the Fies selection process, were out of higher education, according to the Higher Education Census. Except for models (2) and (6), the models were estimated with the choice of the optimal bandwidth based on (Calonico et al., 2014, 2017). Models (5) to (8) are clustered in the standard error at the level of the shift (morning, afternoon, or evening), degree, higher education institution, and selection process, which corresponds to the level of selection of candidates. Robust standard errors in parentheses. The first stage coefficient in these models ranged from 5.2 p.p. to 10.6 p.p., always significant at 1%. "N" considers the number of observations of the sample used, and "n" is the number of observations within the established bandwidth. *** p <0.01, ** p < 0.05, * p < 0.10. OP is the order polynomial.

education in a course with a cheaper semester instead of the preferred degree. With the sample of students who were not in higher education, we analyzed the impact of Fies on

access to courses. The Figure illustrates the effect of Fies on access to degrees. The results in Table 2 are very similar to those previously described. The effect of Fies on the impact of access to degrees is positive, statistically significant, and robust. Students with Fies above the final passing grade enter the applied degree with a chance of 36.2 percentage points greater than those below this cutoff score if I consider the semester of application to the program. Up to 2 years later, this impact drops to 24.2 percentage points, still significant at 1%.

Table 2: Effects of Fies on access to non-higher education students' degrees for those outside higher education - final passing grade

						Cluster-robust		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Access in the semester							
	0.597***	0.596***	0.408***	0.362***	0.596***	0.599***	0.411***	0.362***
	(0.029)	(0.025)	(0.049)	(0.056)	(0.031)	(0.035)	(0.05)	(0.057)
Bandwidth	14.288	30	14.781	42.122	14.48	12.133	14.917	42.223
n	166,739	305,142	55,147	112,380	167,408	132,118	55,379	112,578
	Access up to one year after							
	0.538***	0.539***	0.264***	0.256***	0.535***	0.536***	0.268***	0.256***
	(0.034)	(0.026)	(0.048)	(0.052)	(0.035)	(0.039)	(0.049)	(0.052)
Bandwidth	11.248	30	12.687	37.084	11.579	10.587	12.773	37.119
n	137,783	305,142	50,702	102,138	140,026	118,592	50,877	102,183
	Access up to two years after							
	0.537***	0.533***	0.252***	0.242***	0.535***	0.54***	0.252***	0.242***
	(0.034)	(0.026)	(0.048)	(0.051)	(0.036)	(0.04)	(0.048)	(0.052)
Bandwidth	11.127	30	12.088	35.563	11.226	10.529	12.267	35.571
n	136,625	305,142	49,408	98,909	136,729	118,059	49,803	98,909
N	770,002	770,002	230,128	230,128	765,998	693,078	230,128	230,128
OP	1	1	1	2	1	1	1	2
Kernel	Uniform	Uniform	Uniform	Triangular	Uniform	Uniform	Uniform	Triangular
Covariates	N	N	S	S	N	N	S	S

Note: The sample used only considers students who, between 2010 and the year of enrollment in the selection process, were outside the higher education degree for which they enrolled in Fies, according to the Higher Education Census. Except for models (2) and (6), the models were estimated with the choice of the optimal bandwidth based on (Calonico et al., 2014, 2017). Models (5) to (8) are clustered in the standard error at the level of the shift (morning, afternoon, or evening), degree, higher education institution, and selection process, which corresponds to the level of selection of candidates. Robust standard errors in parentheses. The first stage coefficient in these models ranged from 4.4 p.p. to 10.7 p.p., always significant at 1%. "N" considers the number of observations of the sample used, and "n" is the number of observations within the established bandwidth. *** p <0.01, ** p < 0.05, * p < 0.10. OP is the order polynomial.

5.2 Higher-education permanence

Table 3 shows the impacts of Fies on permanence in higher education, that is, the condition of active enrollment throughout the degree in the years following the student enrollment process in the Fies selection process. I am waiting for the results to be sent, after using the Inep secrecy room, to complement the results in this table. The impact of Fies on

enrollment is positive and statistically significant from the year following the Fies selection process to 4 years later. The effect is greater in the following year, with an estimate of around 32 p.p., and drops over the years, with an estimated effect of 13.7 p.p. four years after enrollment in the Fies selection process.

Table 3: Effects of Fies on higher-education permanence (active enrollment condition) - final passing grade

			Cluster-robust	
	(1)	(2)	(3)	(4)
	Enrollment next year			
	0.323***	0.321***	0.323***	0.32***
	(0.032)	(0.024)	(0.033)	(0.025)
Bandwidth	14.546	30	14.559	30
n	267,952	470,084	265,940	466,796
	Enrollment two years after			
	0.233***	0.218***	0.234***	0.218***
	(0.03)	(0.026)	(0.032)	(0.027)
Bandwidth	17.093	30	17.087	30
n	302,581	470,084	300,352	466,796
	Enrollment three years after			
	0.21***	0.209***	0.212***	0.21***
	(0.028)	(0.027)	(0.03)	(0.028)
Bandwidth	20.06	30	19.767	30
n	342,705	470,084	336,171	466,796
	Enrollment four years after			
	0.154***	0.136***	0.154***	0.137***
	(0.036)	(0.026)	(0.038)	(0.028)
Bandwidth	13.964	30	13.773	30
	259,925	470,084	255,258	466,796
N	1,220,484	1,220,484	1,214,213	1,214,213
OP	1	1	1	1
Kernel	Uniform	Uniform	Uniform	Uniform
Covariates	N	N	N	N

Note: The sample used considers all students enrolled in the Fies selection process. The models (2) and (3) were estimated with the choice of the optimal bandwidth based on (Calonico et al., 2014, 2017). Models (3) to (4) are clustered in the standard error at the level of the shift (morning, afternoon, or evening), degree, higher education institution, and selection process, which corresponds to the level of selection of candidates. Robust standard errors in parentheses. The first stage coefficient in these models ranged from 1.3 p.p. to 10.9 p.p., always significant at 1%. "N" considers the number of observations of the sample used, and "n" is the number of observations within the established bandwidth. *** p <0.01, ** p < 0.05, * p < 0.10. OP is the order polynomial.

5.3 Degree migration

The initial results of Table A.3 show that the effect on the degree migration is null because when I modify the RDD estimation bandwidth, there is a sign inversion without statistical

significance. Thus, it can be interpreted that Fies does not encourage students to change courses while they are in higher education. This table will be complemented with the results of the Inep waiting extraction, which will include the estimates with control and change in the order of the polynomial and the used kernel.

Table 4: Effects of Fies on degree migration during higher education

			Cluster-robust	
	(1)	(2)	(3)	(4)
(5)	(6)	(7)	(8)	
	-0.003	0.005	-0.003	0.005
	(0.005)	(0.006)	(0.005)	(0.006)
Bandwidth	34.269	30	33.874	30
n	341397	307183	335390	304571
N	770,839	770,839	766,043	766,043
OP	1	1	1	1
Kernel	Uniform	Uniform	Uniform	Uniform
Covariates	N	N	N	N

Note: The sample used considers all students enrolled in the Fies selection process. The models (2) and (3) were estimated with the choice of the optimal bandwidth based on (Calonico et al., 2014, 2017). Models (3) to (4) are clustered in the standard error at the level of the shift (morning, afternoon, or evening), degree, higher education institution, and selection process, which corresponds to the level of selection of candidates. Robust standard errors in parentheses. The first stage coefficient in these models ranged from 10.9 p.p. to 11.3 p.p., always significant at 1%. "N" considers the number of observations of the sample used, and "n" is the number of observations within the established bandwidth. *** p <0.01, ** p < 0.05, * p < 0.10. OP is the order polynomial.

5.4 Degree completion time

It is important to consider that the results on completion consider a smaller sample because many students who entered the analyzed period are still in degrees with expected completion forecast for 2022. For example, those who entered in the second half of 2017 are in a degree of 5 years. My sample considers students who entered and could conclude by 2021, the last year of our database, in the graduate result variable in the expected year. In the variables graduated up to one year or two years later, the expected year of completion is up to, respectively, 2020 and 2019. Current results show that Fies students above the passing grade have a lower chance of around 33.5 p.p. to complete the degree in the expected year. Regarding the chance of graduating in up to one or two more years, waiting for the results with controls for a more conclusive analysis is necessary.

6 Robustness checks

Key robustness results will be presented with RD donut designs, with variations in donut holes, followingCattaneo et al. (2019). These results have already been run and will be

Table 5: Effects of Fies on degree completion time

			Cluster robust	
	(1)	(2)	(3)	(4)
	Graduates in expected year			
	-0.355***	-0.344***	-0.346***	-0.335***
	(0.047)	(0.033)	(0.047)	(0.034)
Bandwidth	12.314	30	12.385	30
n	143,974	290,978	143,267	288,366
N	725364	725364	720568	720568
	Graduates one year later			
	0.076*	0.064	0.075*	0.064
	(0.04)	(0.044)	(0.041)	(0.045)
Bandwidth	28.563	30	28.824	30
n	77447	80294	77979	80294
N	200752	200752	200752	200752
	Graduates two year later			
	0.065*	0.011	0.065*	0.011
	(0.036)	(0.056)	(0.036)	(0.058)
Bandwidth	45.643	30	45.543	30
n	39935	29596	39881	29596
N	68846	68846	68846	68846
OP	1	1	1	2
Kernel	Uniform	Uniform	Uniform	Triangular
Covariates	N	N	S	S

Note: The sample used considers all students enrolled in the Fies selection process. The models (2) and (3) were estimated with the choice of the optimal bandwidth based on (Calonico et al., 2014, 2017). Models (3) to (4) are clustered in the standard error at the level of the shift (morning, afternoon, or evening), degree, higher education institution, and selection process, which corresponds to the level of selection of candidates. Robust standard errors in parentheses. The first stage coefficient in these models ranged from 8.1 p.p. to 12.3 p.p., always significant at 1%. "N" considers the number of observations of the sample used, and "n" is the number of observations within the established bandwidth. *** p <0.01, ** p < 0.05, * p < 0.10. OP is the order polynomial.

received next month, with the extraction of Inep. As an example, Table 6 shows the results of the RD donut estimates for access to higher education, removing the relative score from -3 to 3 from the estimates. These estimates show that the impact of Fies on access to higher education is always significant at 1% and with a coefficient even higher than those obtained in the main results. In addition to these robustness tests, the appendix presents the results with a sample excluding the second semester of 2015, as it was the first time that the government applied the rule in that semester, its use was flexible, and the results considering the passing grade for the first call, in which the student selection occurs yet without the use of waiting list. The results available in the appendix show that the impact of Fies on access to higher education and degrees of interest is robust.

Table 6: Donut RD (3 points) - Access to higher education for those outside higher education - final passing grade

	(1)	(2)	(3)	(4)	(5)	(6)	Cluster-robust (7)	(8)
				Access in the semester				
	0.605***	0.623***	0.605***	0.623***	0.609***	0.62***	0.602***	0.623***
	(0.023)	(0.039)	(0.014)	(0.021)	(0.025)	(0.039)	(0.014)	(0.021)
Bandwidth	26.296	30	38.93	57.297	25.539	30	39.602	57.39
n	238,344	272,544	84,810	124,234	229,601	270,755	86,382	124,461
				Access up to one year after				
	0.433***	0.474***	0.33***	0.323***	0.435***	0.468***	0.329***	0.323***
	(0.025)	(0.046)	(0.014)	(0.019)	(0.027)	(0.047)	(0.014)	(0.019)
Bandwidth	31.274	30	41.167	72.588	30.622	30	40.569	72.032
n	283,908	272,544	89,925	152,149	276,434	270,755	88,561	151,236
				Access up to two years after				
	0.425***	0.473***	0.313***	0.311***	0.431***	0.467***	0.313***	0.311***
	(0.025)	(0.047)	(0.013)	(0.019)	(0.027)	(0.047)	(0.014)	(0.019)
Bandwidth	31.241	30	41.79	71.165	30.209	30	41.2	70.546
n	283,761	272,544	91,268	149,798	272,670	270,755	89,967	148,698
N	781,543	781,543	222,380	222,380	777,792	777,792	222,380	222,380
OP	1	1	1	2	1	1	1	2
Kernel	Uniform	Uniform	Uniform	Triangular	Uniform	Uniform	Uniform	Triangular
Covariates	N	N	S	S	N	N	S	S

Note: The sample used only considers students who, between 2010 and the year of enrollment in the Fies selection process, were out of higher education, according to the Higher Education Census. Except for models (2) and (6), the models were estimated with the choice of the optimal bandwidth based on (Calonico et al., 2014, 2017). Models (5) to (8) are clustered in the standard error at the level of the shift (morning, afternoon, or evening), degree, higher education institution, and selection process, which corresponds to the level of selection of candidates. Robust standard errors in parentheses. The first stage coefficient in these models ranged from 17.4 p.p. to 61.9 p.p. (models with triangular kernel and polynomial of order 2), always significant at 1%. "N" considers the number of observations of the sample used, and "n" is the number of observations within the established bandwidth. The donut hole is 3 points each side, so the -3 to 3 range is outside the estimations. *** p <0.01, ** p < 0.05, * p < 0.10. OP is the order polynomial.

7 Conclusion

Using the Fies' administrative data, which contains those applied in the selection processes and the management of signed contracts, and the change made in the program selection process in 2015, which became competitive in degrees with greater demand than supply and, as a result, established a passing grade for these degrees, it was possible to evaluate through of the discontinuity regression designs and with greater internal validity the impacts of student loan on access, permanence, migration, and graduation in higher education. The results of this work show that there is a positive and significant impact of Fies on access and permanence. On the other hand, the chance of the Fies student completing the degree in the expected time is lower. Fies does not change students' incentives about course migration once they are in higher education. The results shown here will complement robustness tests already conducted in the Inep secrecy room used

for the research development.

This work makes a significant contribution by assessing the effects of student loans on higher education access. Before this study, there were no robust impact assessments on this matter, as existing databases only covered enrolled students, not those interested but not yet enrolled. I evaluate the effects of access to higher education in the semester of enrollment in the Fies selection process, in the following year, and up to 2 years later. The analysis of the following years assumes that students who didn't receive the student loan might take a little longer to enter higher education. Considering the students who, from 2010 until the year of registration in the Fies selection process, were out of higher education, I show that access to student loans, in the analyzed design, increases the fraction of students with a score above the cut-off point who access the higher education by 51.2 percentage points, in the program registration semester. This high impact is maintained a year or two later, equivalent to 31.5 percentage points in both years. One question is whether students are accessing the course they signed up for in Fies, so I analyze the impact of Fies on access to these degrees. Fies increases the fraction of students with a score above the passing grade who access the applied degree by 41.1 percentage points in the semester of the selection process. The fraction of students with scores above the passing grade and with access to student loans in the applied degree one year and two years after the selection process is, respectively, 26.8 and 25.2 percentage points higher than the students below the passing grade.

Bibliography

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2020). Sampling-based versus design-based uncertainty in regression analysis. *Econometrica*, 88(1):265–296.

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, 138(1):1–35.

Bailey, M. and Dynarski, S. M. (2011). Inequality in postsecondary attainment.

Becker, G. S. (1967). *Human capital and the personal distribution of income: An analytical approach*. Number 1. Institute of Public Administration.

Becker, K. L. and Mendonça, M. J. C. d. (2021). Student funding policies: an analysis of the impact of fies on the completion time of higher education. *Economia e Sociedade*, 30:551–581.

Belley, P., Frenette, M., and Lochner, L. (2010). Post-secondary attendance by parental income: Comparing the us and canada. Technical report, CIBC Working Paper.

Belley, P. and Lochner, L. (2007). The changing role of family income and ability in determining educational achievement. *Journal of Human capital*, 1(1):37–89.

Brown, M., Karl Scholz, J., and Seshadri, A. (2012). A new test of borrowing constraints for education. *The Review of Economic Studies*, 79(2):511–538.

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

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, 17(2):372–404.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Cameron, S. V. and Heckman, J. J. (2001). The dynamics of educational attainment for black, hispanic, and white males. *Journal of political Economy*, 109(3):455–499.

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics*, 3:1801–1863.

Card, D. and Solis, A. (2022). Measuring the effect of student loans on college persistence. *Education Finance and Policy*, 17(2):335–366.

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

Carver, A. B. et al. (2007). The market failure for student loans. *Journal of Private Enterprise*, 23(1):101–115.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2019). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press.

De Mello, J. M. and Duarte, I. F. (2020). The effect of the availability of student credit on tuition. *Economía*, 20(2):179–222.

Dente, B. and Piraino, N. (2011). Models for determining the efficiency of student loans policies. *Journal of Higher Education Policy and Management*, 33(4):375–386.

Dobbin, C., Barahona, N., and Otero, S. (2021). The equilibrium effects of subsidized student loan. Technical report, Working paper, Stanford University.

Dynarski, S. M. (2003). Does aid matter? measuring the effect of student aid on college attendance and completion. *American Economic Review*, 93(1):279–288.

Estevan, F., Gall, T., and Morin, L.-P. (2019). Redistribution without distortion: Evidence from an affirmative action programme at a large brazilian university. *The Economic Journal*, 129(619):1182–1220.

Fack, G. and Grenet, J. (2015). Improving college access and success for low-income

students: Evidence from a large need-based grant program. *American Economic Journal: Applied Economics*, 7(2):1–34.

Francis, A. M. and Tannuri-Pianto, M. (2012). The redistributive equity of affirmative action: Exploring the role of race, socioeconomic status, and gender in college admissions. *Economics of Education Review*, 31(1):45–55.

Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3):933–959.

Kane, T. J. (1996). College cost, borrowing constraints and the timing of college entry. *Eastern Economic Journal*, 22(2):181–194.

Katz, L. F. et al. (1999). Changes in the wage structure and earnings inequality. In *Handbook of labor economics*, volume 3, pages 1463–1555. Elsevier.

Korting, C., Lieberman, C., Matsudaira, J., Pei, Z., and Shen, Y. (2021). Visual inference and graphical representation in regression discontinuity designs. *arXiv preprint arXiv:2112.03096*.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

Liu, Y., Green, A., and Pensiero, N. (2016). Expansion of higher education and inequality of opportunities: a cross-national analysis. *Journal of Higher Education Policy and Management*, 38(3):242–263.

Lochner, L. and Monge-Naranjo, A. (2016). Student loans and repayment: Theory, evidence, and policy. In *Handbook of the Economics of Education*, volume 5, pages 397–478. Elsevier.

Lochner, L. J. and Monge-Naranjo, A. (2011). The nature of credit constraints and human capital. *American economic review*, 101(6):2487–2529.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

Mello, U. (2023). Affirmative action and the choice of schools. *Journal of Public Economics*, 219:104824.

Montoya, A. M., Noton Norambuena, C., and Solis, A. (2018). The returns to college choice: Loans, scholarships and labor outcomes. Technical report, Working Paper.

Palmisano, F., Biagi, F., and Peragine, V. (2022). Inequality of opportunity in tertiary education: evidence from europe. *Research in Higher Education*, pages 1–52.

Pontuschka, R. (2016). *Avaliação de impacto do Fies*. PhD thesis, Universidade de São Paulo.

Rocha, W. M., Ehrl, P., Monasterio, L. M., et al. (2016). Análise de impacto do fies sobre o salário do trabalhador formal. *44º Encontro Nacional de Economia-Anpec*.

Sewell, W. H. (1971). Inequality of opportunity for higher education. *American Sociological Review*, 36(5):793–809.

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

Sun, S. T. and Yannelis, C. (2016). Credit constraints and demand for higher education: Evidence from financial deregulation. *Review of Economics and Statistics*, 98(1):12–24.

Valente, R. R. and Berry, B. J. (2017). Performance of students admitted through affirmative action in brazil. *Latin American Research Review*, 52(1):18–34.

Vieira, R. S. and Arends-Kuenning, M. (2019). Affirmative action in brazilian universities:

Effects on the enrollment of targeted groups. *Economics of Education Review*, 73:101931.

Zimmerman, S. D. (2014). The returns to college admission for academically marginal students. *Journal of Labor Economics*, 32(4):711–754.

A.1 Appendix

Table A.1: Descriptive statistics of Enem applications, Fies applications and first-time Fies applicants

	Enem applicants			Fies applicants		
Variable	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Woman	19,242,666	0.59	0.49	1,838,582	0.61	0.49
Single	19,242,666	0.86	0.35	1,838,582	0.90	0.31
Disabled person	19,242,666	0.00	0.03	1,838,582	0.00	0.01
Black, brown ou indigenous	19,242,666	0.55	0.50	1,838,582	0.59	0.49
Public high school	11,318,017	0.81	0.39	1,838,309	0.63	0.48
Regular high school	19,242,666	0.65	0.48	1,838,582	0.55	0.50
Age	19,242,562	22.84	7.89	1,838,580	22.13	6.39
Uneducated father	19,242,666	0.06	0.25	1,838,582	0.04	0.20
Father with incomplete high school	19,242,666	0.51	0.50	1,838,582	0.49	0.50
Father with complete high school	19,242,666	0.23	0.42	1,838,582	0.29	0.46
Father with higher education	19,242,666	0.07	0.25	1,838,582	0.07	0.25
Father with a postgraduate	19,242,666	0.03	0.17	1,838,582	0.03	0.16
Uneducated mother	19,242,666	0.05	0.22	1,838,582	0.03	0.17
Mother with incomplete high school	19,242,666	0.49	0.50	1,838,582	0.44	0.50
Mother with complete high school	19,242,666	0.28	0.45	1,838,582	0.35	0.48
Mother with higher education	19,242,666	0.09	0.28	1,838,582	0.10	0.29
Mother with a postgraduate	19,242,666	0.05	0.22	1,838,582	0.06	0.24
Number of residents in the household	19,170,426	4.06	1.63	1,838,289	3.91	1.52
Times applied Enem	19,242,666	1.81	1.09	1,838,582	2.66	1.35
From 0 to 5, how much Fies reason for Enem	15,909,601	3.59	1.82	1,808,852	4.13	1.44
Dummy for interest in Fies	13,100,910	0.76	0.43	1,327,346	0.86	0.35
Average score Enem on best performance	18,285,038	519.56	78.55	1,836,751	538.63	59.53
Essay score on best performance	19,242,666	525.70	187.20	1,838,582	596.12	123.21
Human sciences score*	19,199,480	531.67	82.42	1,838,466	555.98	62.81
Natural sciences score*	19,199,480	475.62	76.23	1,838,466	486.31	69.02
Mathematics score*	18,805,263	490.18	111.32	1,837,539	491.52	98.77
Reading and codes score*	18,805,263	508.28	75.00	1,837,539	528.07	58.73
Essays score*	19,242,666	494.19	195.45	1,838,582	561.84	139.37

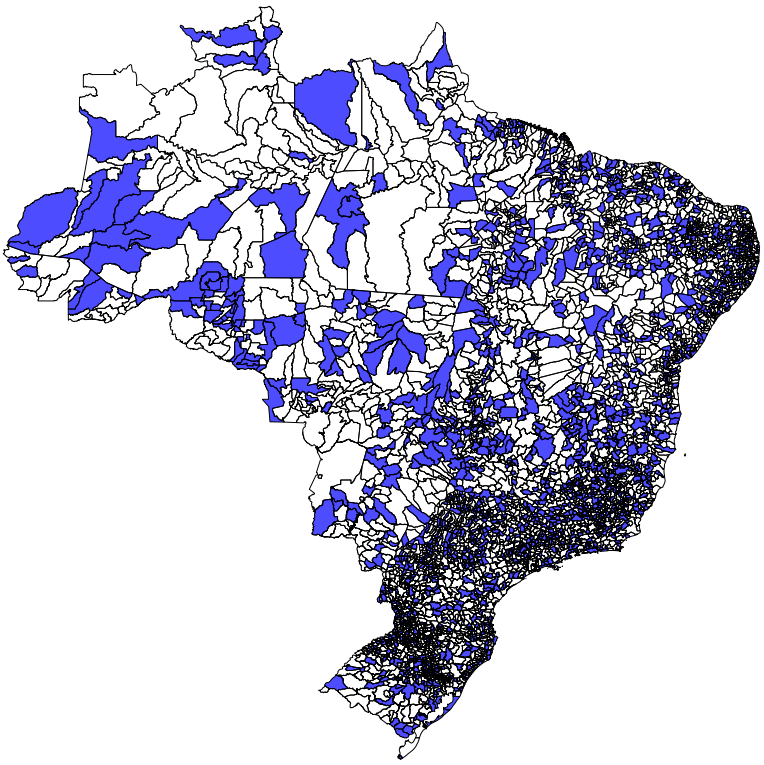
Source: Enem database and administrative database of the program's selection processes.
*Scores from the last time the student applied the Enem.

Table A.2: Discontinuity Regression for the covariates

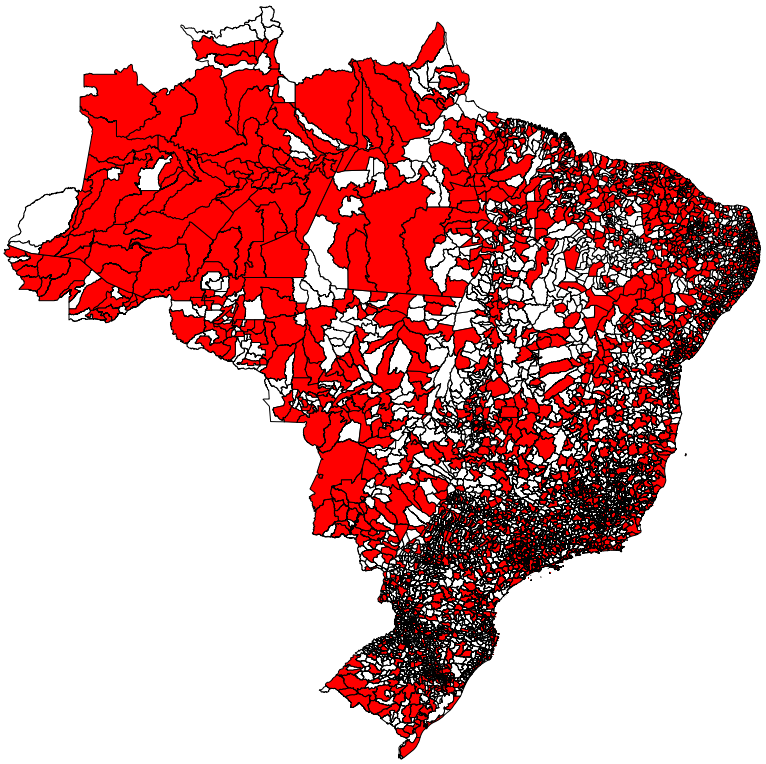
Woman	RD Effect	St. Dev.	Bandwidth optimal	n	N
Single	-0.034	0.062	22.137	168286	425124
Age					
Black, brown and indigenous					
Disable person	0.002**	0.001	31.546	529953	1328930
Public high school	0.116	0.151	14.051	286254	1327601
Regular education	0.099	0.126	15.263	303338	1328930
Uneducated father	-0.065	0.047	20.14	374061	1328930
Father with incomplete high school	0.073	0.161	13.485	277065	1328930
Father with high school	0.198	0.12	16.575	322329	1328930
Father with higher education	0.079	0.066	16.58	322618	1328930
Father with a postgraduate degree	-0.103***	0.036	22.366	405510	1328930
Uneducated mother	-0.012	0.035	21.96	399900	1328930
Mother with incomplete high school	-0.06	0.144	14.467	291522	1328930
Mother with high school	-0.078	0.137	14.533	292427	1328930
Mother with higher education	-0.015	0.075	17.646	338083	1328930
Mother with a postgraduate degree	0.083	0.061	18.779	354453	1328930
# of residents	0.357	0.493	13.411	275803	1328714
Times applied Enem	0.413	0.45	13.204	272848	1328930
Times aplied Fies	-2.431***	0.354	12.87	271060	1345025
Technological degree	7.642**	3.024	12.175	167620	806003
Fies reason for Enem	-0.388	0.443	13.681	275472	1307460
Reported interest in Fies in Enem	0.079	0.144	13.248	197434	959477
Score in the essay in the highest average	10.357	36.645	13.754	280644	1327626
Discounted semester tuition fee	-16601.803***	983.331	21.567	373945	1281532
Per capita family income	1443.801***	271.678	13.218	276197	1345020
Competition for vacancy	2.235**	1.128	12.22	261337	1344833

Source: Enem database, administrative database of the program's selection processes, Student contract management database in the program funding phase and Higher Education Census.
Notes: Estimates obtained with package from (Calonico et al., 2017), with the choice of optimal bandwith, polynomial of degree 1 and uniform kernel. Scores regarding final passing grade.

Figure A.1: Municipalities with application of the Enem in the years 2010 and 2016



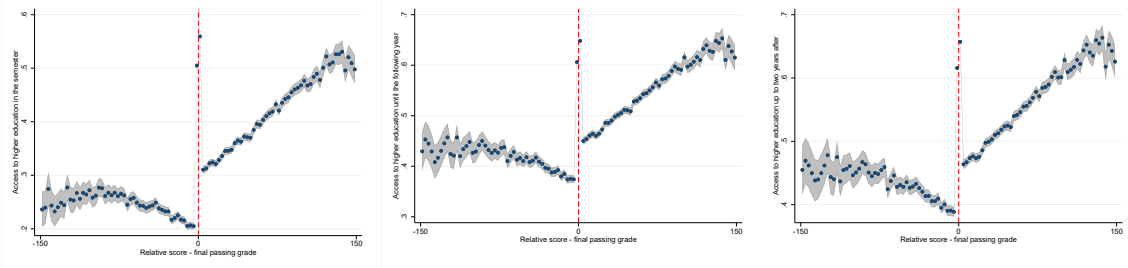
(a) In 2010



(b) In 2016

Source: Enem Administrative Records.

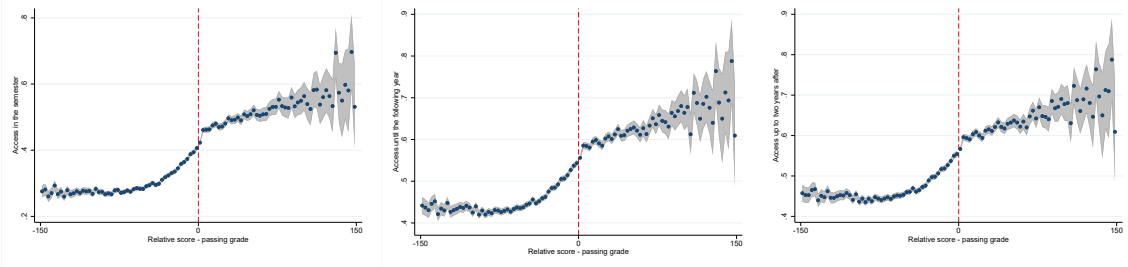
Figure A.2: RD Plot - Access to higher education with final passing grade



(a) In the semester (b) Up to one year after (c) Up to two years after

Source: Administrative database of the Fies selection processes, Student contract management database in the program funding phase and Higher Education Census.
Note: Vertical line represents the cutoff score, where the Enem score on the horizontal axis is about the cutoff score given by the last call of the program's selection process. Each dot represents the number of students in a 3-point bin of Enem scores. The shaded region shows 95% confidence intervals.

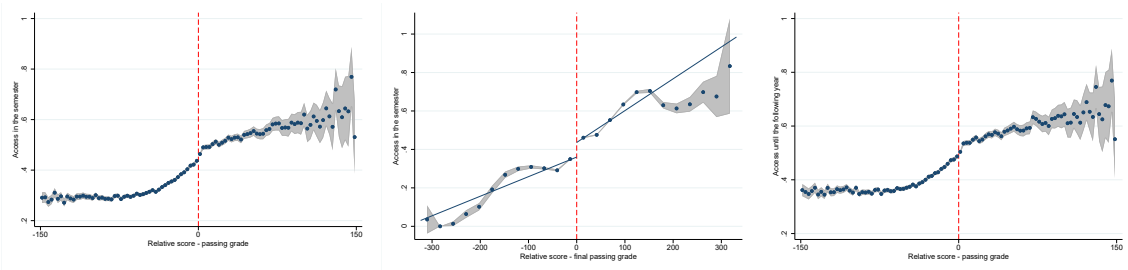
Figure A.3: RD Plot - Access to higher education with first passing grade



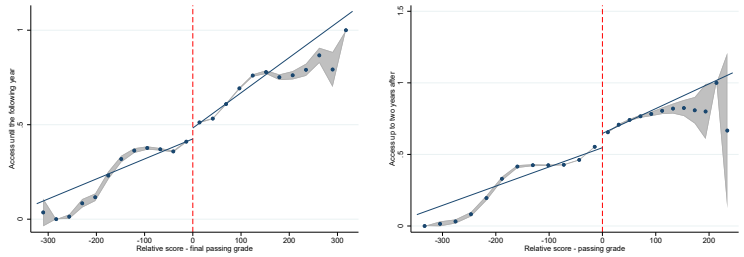
(a) In the semester (b) Up to one year after (c) Up to two years after

Source: Administrative database of the Fies selection processes, Student contract management database in the program funding phase and Higher Education Census.
Note: Vertical line represents the cutoff score, where the Enem score on the horizontal axis is about the cutoff score given by the last call of the program's selection process. Estimates obtained with package from (Calonico et al., 2017), with the polynomial of degree 1 and uniform kernel. Each dot represents the number of students in a 3-point bin of Enem scores. The shaded region shows 95% confidence intervals.

Figure A.4: RD Plot - Access to non-higher education students' degrees



(a) In the semester - first passing grade (b) In the semester - final passing grade (c) Up to one year after - first passing grade

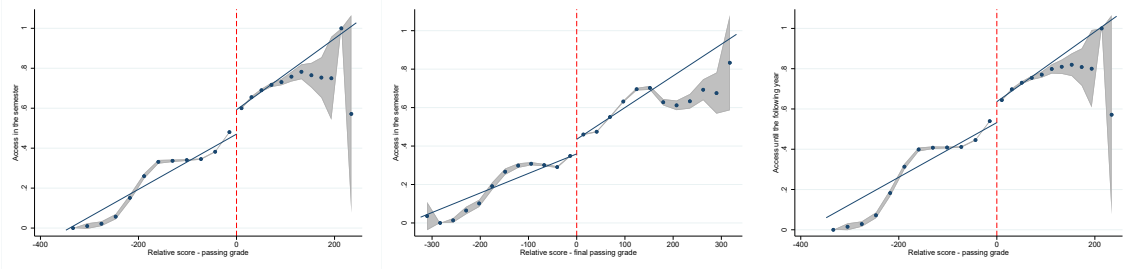


Cap1/Figuras/acesso_curso

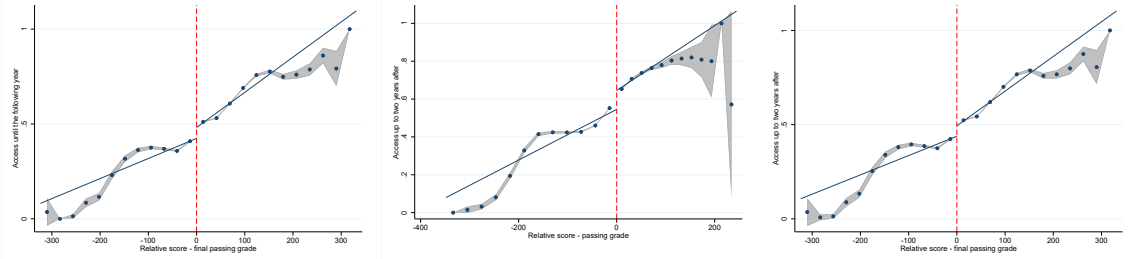
(d) Up to one year after - final passing grade (e) Up to two years after - first passing grade (f) Up to two years after - final passing grade

Source: Selection Fies, Fies Computerized System and Higher Education Census.
Note: Vertical line represents the cutoff score, where the Enem score on the horizontal axis is about the cutoff score given by the first or last call of the program's selection process. The solid line represents the fitted values of the polynomial of degree 1 with uniform kernel. The graph shows optimal bins. The shaded region shows 95% confidence intervals.

Figure A.5: RD Plot - Degree access for off-degree enrollees



(a) In the semester - first passing grade (b) In the semester - final passing grade (c) Up to one year after - first passing grade



(d) Up to one year after - final passing grade (e) Up to two years after - first passing grade (f) Up to two years after - final passing grade

Source: Selection Fies, Fies Computerized System and Higher Education Census.
Note: Vertical line represents the cutoff score, where the Enem score on the horizontal axis is about the cutoff score given by the first or last call of the program's selection process. The solid line represents the fitted values of the polynomial of degree 1 with uniform kernel. The graph shows optimal bins. The shaded region shows 95% confidence intervals.

Table A.3: Access to higher education for those outside higher education without the 2015 selection process - final passing grade

				Cluster-robust				
	(1)	(2)	(3)	(4)				
				Access in the semester				
	0.572***	0.584***	0.522***	0.502***	0.568***	0.582***	0.514***	0.502***
	(0.035)	(0.028)	(0.068)	(0.088)	(0.037)	(0.03)	(0.069)	(0.088)
Bandwidth	12.941	30	15.179	39.906	13.106	30	15.102	40.065
n	153,579	304,390	55,435	107,387	154,124	302,362	55,312	107,708
				Access until the following year				
	0.432***	0.443***	0.357***	0.36***	0.43***	0.438***	0.357***	0.36***
	(0.042)	(0.032)	(0.067)	(0.087)	(0.043)	(0.033)	(0.067)	(0.087)
Bandwidth	11.815	30	15.368	39.479	11.803	30	15.357	39.497
n	142,630	304,390	55,853	106,452	141,768	302,362	55,797	106,503
				Access up to two years after				
	0.431***	0.442***	0.345***	0.351***	0.43***	0.437***	0.345***	0.351***
	(0.042)	(0.032)	(0.066)	(0.086)	(0.043)	(0.033)	(0.067)	(0.086)
Bandwidth	11.459	30	15.643	39.512	11.523	30	15.636	39.542
n	139,154	304,390	56,409	106,540	139,075	302,362	56,365	106,622
N	765,217	765,217	229,301	229,301	761,245	761,245	229,301	229,301
OP	1	1	1	2	1	1	1	2
Kernel	Uniform	Uniform	Uniform	Triangular	Uniform	Uniform	Uniform	Triangular
Covariates	N	N	S	S	N	N	S	S

Note: The sample used only considers students who, between 2010 and the year of enrollment in the Fies selection process, were out of higher education, according to the Higher Education Census. Except for models (2) and (6), the models were estimated with the choice of the optimal bandwidth based on (Calonico et al., 2014, 2017). Models (5) to (8) are clustered in the standard error at the level of the shift (morning, afternoon, or evening), degree, higher education institution, and selection process, which corresponds to the level of selection of candidates. Robust standard errors in parentheses. The first stage coefficient in these models ranged from 5.2 p.p. to 10.6 p.p., always significant at 1%. "N" considers the number of observations of the sample used, and "n" is the number of observations within the established bandwidth. *** p <0.01, ** p < 0.05, * p < 0.10. OP is the order polynomial.

Table A.4: Access to higher education for those outside higher education - first passing grade

						Cluster-robust		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Access in the semester				
	1.413***	1.207***	0.657***	0.597***	1.392***	1.252***	0.65***	0.597***
	(0.355)	(0.129)	(0.082)	(0.073)	(0.283)	(0.125)	(0.082)	(0.076)
Bandwidth	15.072	30	24.804	75.887	14.898	30	25.189	75.324
n	194,494	332,858	110,384	213,643	190,740	329,796	111,526	213,028
				Access until the following year				
	1.131*	0.894***	0.339***	0.282***	0.992***	0.909***	0.342***	0.282***
	(0.511)	(0.153)	(0.08)	(0.075)	(0.345)	(0.145)	(0.082)	(0.079)
Bandwidth	12.741	30	25.351	76.903	12.925	30	25.598	75.494
n	171,703	332,858	112,015	214,790	171,660	329,796	112,758	213,203
				Access up to two years after				
	0.914*	0.859***	0.358***	0.296***	0.931***	0.884***	0.384***	0.295***
	(0.483)	(0.154)	(0.08)	(0.075)	(0.344)	(0.146)	(0.086)	(0.079)
Bandwidth	12.587	30	23.982	75.233	12.952	30	24.778	73.978
n	170,032	332,858	107,784	212,916	171,870	329,796	110,256	211,513
N	865,154	865,154	251,377	251,377	858,924	858,924	251,377	251,377
OP	1	1	1	2	1	1	1	2
Kernel	Uniform	Uniform	Uniform	Triangular	Uniform	Uniform	Uniform	Triangular
Covariates	N	N	S	S	N	N	S	S

Note: The sample used only considers students who, between 2010 and the year of enrollment in the Fies selection process, were out of higher education, according to the Higher Education Census. Except for models (2) and (6), the models were estimated with the choice of the optimal bandwidth based on (Calonico et al., 2014, 2017). Models (5) to (8) are clustered in the standard error at the level of the shift (morning, afternoon, or evening), degree, higher education institution, and selection process, which corresponds to the level of selection of candidates. Robust standard errors in parentheses. The first stage coefficient in these models ranged from 5.2 p.p. to 10.6 p.p., always significant at 1%. "N" considers the number of observations of the sample used, and "n" is the number of observations within the established bandwidth. *** p <0.01, ** p < 0.05, * p < 0.10. OP is the order polynomial.

Table A.5: Access to non-higher education students' degrees without the 2015 selection process - final passing grade

						Cluster-robust		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Access in the semester	0.607***	0.613***	0.417***	0.361***	0.599***	0.61***	0.426***	0.362***
	(0.033)	(0.026)	(0.063)	(0.071)	(0.035)	(0.028)	(0.063)	(0.072)
Bandwith	11.822	30	10.906	32.228	12.133	30	11.086	32.486
n	130,387	277,784	41,412	81,236	132,118	275,756	41,775	81,713
Access until the following year	0.54***	0.543***	0.307***	0.257***	0.536***	0.539***	0.308***	0.257***
	(0.037)	(0.028)	(0.055)	(0.066)	(0.039)	(0.029)	(0.055)	(0.067)
Bandwith	10.506	30	11.822	30.388	10.587	30	11.911	30.423
n	118,672	277,784	43,152	77,859	118,592	275,756	43,287	77,933
Access up to two years after	0.538***	0.54***	0.277***	0.239***	0.54***	0.536***	0.277***	0.24***
	(0.038)	(0.028)	(0.054)	(0.064)	(0.04)	(0.03)	(0.055)	(0.065)
Bandwith	10.427	30	11.387	30.629	10.529	30	11.483	30.704
n	117,895	277,784	42,349	78,294	118,059	275,756	42,522	78,438
N	697,050	697,050	204,605	204,605	693,078	693,078	204,605	204,605
Order polynomial	1	1	1	2	1	1	1	2
Kernel	Uniform	Uniform	Uniform	Triangular	Uniform	Uniform	Uniform	Triangular
Covariates	N	N	S	S	N	N	S	S

Table A.6: Access to non-higher education students' degrees - first passing grade

						Cluster-robust		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Access in the semester	1.449***	1.21***	0.592***	0.582***	1.416***	1.27***	0.59***	0.583***
	(0.313)	(0.123)	(0.069)	(0.068)	(0.242)	(0.118)	(0.07)	(0.068)
Bandwith	15.231	30	25.287	59.576	16.074	30	23.739	59.012
n	178,737	303,424	101,100	172,706	184,001	300,362	96,735	171,909
Access until the following year	0.862***	0.783***	0.383***	0.363***	1.062***	0.933***	0.403***	0.367***
	(0.314)	(0.133)	(0.072)	(0.069)	(0.26)	(0.126)	(0.069)	(0.07)
Bandwith	13.77	30	20.078	62.965	13.812	30	20.797	62.113
n	165,739	303,424	85,737	177,375	164,204	300,362	87,992	176,257
Access up to two years after	0.74**	0.734***	0.409***	0.379***	0.986***	0.897***	0.411***	0.382***
	(0.343)	(0.135)	(0.067)	(0.068)	(0.271)	(0.129)	(0.068)	(0.068)
Bandwith	13.77	30	20.078	62.965	13.812	30	20.797	62.113
n	160,562	303,424	90,085	175,151	159,237	300,362	89,283	173,768
N	790,679	790,679	226,018	226,018	784,449	784,449	226,018	226,018
Order polynomial	1	1	1	2	1	1	1	2
Kernel	Uniform	Uniform	Uniform	Triangular	Uniform	Uniform	Uniform	Triangular
Covariates	N	N	S	S	N	N	S	S

Table A.7: Higher-education permanence (active enrollment condition) without the 2015 selection process - final passing grade

						Cluster-robust		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Enrollment next year	0.365***	0.36***	0.423	0.942	0.369***	0.358***	0.388	0.927
	(0.047)	(0.028)	(0.278)	(0.604)	(0.047)	(0.029)	(0.285)	(0.612)
Bandwith	10.28	30	40.994	82.265	10.427	30	39.898	81.294
n	177,474	405,441	127,500	205,687	177,575	402,219	124,109	202,747
Enrollment two years after	0.217***	0.233***	0.684	0.923	0.222***	0.234***	0.687	0.972
	(0.049)	(0.03)	(0.277)	(0.494)	(0.051)	(0.031)	(0.286)	(0.518)
Bandwith	10.46	30	46.035	111.371	10.432	30	44.649	106.803
n	179,673	405,441	138,205	240,528	177,575	402,219	134,146	234,713
Enrollment three years after	0.226***	0.224***	1.536***	2.453***	0.231***	0.225***	1.584***	2.569***
	(0.052)	(0.031)	(0.235)	(0.451)	(0.053)	(0.032)	(0.241)	(0.473)
Bandwith	10.124	30	58.109	135.235	10.28	30	57.981	132.951
n	175,403	405,441	162,829	256,168	175,610	402,219	161,264	253,498
Enrollment four years after	0.132***	0.125***	1.187***	2.575***	0.143***	0.125***	1.254***	2.641***
	(0.047)	(0.03)	(0.22)	(0.579)	(0.048)	(0.032)	(0.23)	(0.609)
Bandwith	11.579	30	62.12	106.61	12.001	30	60.059	103.506
	193,371	405,441	170,542	236,066	197,079	402,219	165,259	231,455
N	1,048,240	1,048,240	270,188	270,188	1,042,073	1,042,073	268,588	268,588
Order polynomial	1	1	1	2	1	1	1	2
Kernel	Uniform	Uniform	Uniform	Triangular	Uniform	Uniform	Uniform	Triangular
Covariates	N	N	S	S	N	N	S	S